Natural Product Biosynthesis in the Genus *Herpetosiphon* – a Genomics and Chemistry-Guided Evaluation

Dissertation

To Fulfill the

Requirements for the Degree of

„doctor rerum naturalium“ (Dr. rer. nat.)

Submitted to the Council of the Faculty
of Biological Sciences
of the Friedrich Schiller University Jena

by Xinli Pan (M.Sc. Chemistry)

born on 11.06.1987 in Nanning, China

Die vorliegende Arbeit wurde am Leibniz-Institut für Naturstoff-Forschung und Infektionsbiologie e.V., Hans-Knöll-Institut Jena, in der Nachwuchsgruppe „Sekundärmetabolismus räuberischer Bakterien“, sowie an der Technischen Universität Dortmund in der Arbeitsgruppe „Technische Biologie“ unter Leitung von Prof. Dr. Markus Nett angefertigt.

Zeitraum: 10. 2013 bis 10. 2017

Gutachter 1: Prof. Dr. Dirk Hoffmeister, Friedrich-Schiller-Universität Jena

Gutachter 2: Prof. Dr. Markus Nett, Technischen Universität Dortmund

Gutachter extern: Prof. Dr. Timo Niedermeyer, Martin-Luther-Universität Halle-Wittenberg

Tag der Verteidigung: 06. 06. 2018

Table of contents

Acknowledgment

I would first like to thank my supervisor, Prof. Dr. M. Nett, who offered me a unique opportunity to work in his lab. He provided invaluable support and guidance on this project. All my words fail to express my gratitude towards my supervisor for his patience, enrich knowledge, motivation, and continuous interest in supervising me.

I would like to thank the China Scholarship Council (CSC) for a doctoral fellowship, which allowed me to study in Germany and to work with plenty of excellent scientists.

I am thankful to my colleagues Sebastian Schieferdecker, Martin Kreutzer and Nicole Domin, who provided valued comments and suggestions. I am especially grateful to Nicole Domin who taught and guided me at the beginning of this project. My sincere thanks also go to Juliane Korp, Dr. Hirokazu Kage, Angela Sester, and Anna Tippelt for the stimulating discussion and for their precious support.

I would like to express my sincere thanks to Heike Heinecke, Andrea Perner, and Karin Martin, who devoted their time and knowledge in the implementation of my project.

I express my gratitude toward my parents and husband for their encouragement and support during this project. Last but not the least, I would like to thank all my friends for supporting and helping me.

1 Introduction

1.1 Need for new antibiotics

Over the centuries, natural products from bacteria, fungi, plants and animals have played a crucial role in drug discovery [1]. The well-known drug penicillin, discovered from the fungus *Penicillium notatum* in 1928, is one of the oldest and most commonly used antibiotics [2]. A number of natural products isolated from plants and fungi have contributed to the development of drugs for the treatment of various diseases. Morphine, which is produced by the plant *Papaver somniferum* L. [3], was first reported in 1803 and is widely used as an analgesic. In 1820, the antimalarial drug quinine was isolated from the cinchona tree [4]. During the golden age of drug discovery (1940s-1960s) [5], several antibiotics were isolated from bacteria. Examples include streptomycin [6], streptothricin [7-8], kanamycin [9], and tetracycline [10] . These drugs are now in medical use for more than 50 years.

Although great efforts have been invested in the discovery of novel antibiotics, mainly known natural product classes have been identified after the golden age. Furthermore, natural product isolation faces several difficulties including low production levels, time-consuming purification and chemical characterization [11-12]. Due to the rapid emergence of antibiotic-resistant bacteria in the entire world, the demand for novel antibiotics is increasing [13]. Several factors contributed to this development, *e.g.*, the heavy use of antibiotics in agriculture and the low degree of biodegradability owing to their chemical nature [14-15]. Typically, bacterial resistance originates from genetic alterations, such as mutation or the transfer of resistance genes from other bacterial strains. It has been well documented that antibiotic-resistant bacteria and their resistance genes are constantly circulating in soils, plants, and animals including livestock [16]. Mechanisms of bacterial resistance include mutations in drug targets, enzymatic inactivation of antibiotics, changes in membrane permeability and activation of efflux pumps to transport antibiotics out of the cell [17]. The spread of drug-resistant bacteria poses a serious threat to public health. Infections by *Acinetobacter* and *Pseudomonas* species can cause life-threatening diseases that include pneumonia, meningitis, and sepsis. These pathogens are resistant to several antibiotics, among them polymyxins, tigecycline,

penicillins, and cephalosporins [18-19]. Enterobacteriaceae are resistant to all known carbapenems [20]. The rate of death from methicillin-resistant *Staphylococcus aureus* (MRSA) infection is higher than HIV, Parkinson's disease, and emphysema [21-23]. Moreover, patients infected by multiresistant pathogens have longer recovery times and higher average mortality rates [24]. In order to decrease the spread of drug-resistant bacteria, it is important to promote the rational use of antibiotics. Furthermore, the discovery of new antibiotics is of utmost importance to replenish our therapeutic arsenal against infectious diseases [25].

1.2 Where to find new antibiotics

The most widely used antibiotics in medicine include the β-lactams (penicillins, cephalosporins and carbapenems) [26-27] and macrolides (erythromycin, fidaxomicin, and azithromycin) [28-31]. The majority of antibiotics are natural products or (semi-)synthetic derivatives of natural products. Examples are amoxicillin [32], cefuroxime [33], azithromycin [34], and clarithromycin [35]. Only the fluoroquinolones [36] and the oxazolidinones (*e.g.*, linezolid) [37] are truly synthetic. Although chemical synthesis often enabled an improvement of the biological activity of naturally occurring antibiotics by introducing structural modifications, it has been of limited use to find new lead structures and to reduce the threat from resistant bacteria [38]. Thus, natural products still represent the most promising source for the discovery of new anti-infective drugs.

Nearly 80% of all antibiotics were discovered from *Streptomyces* species in the 1950s and 1960s [39]. Since then, however, the rate of rediscovery of known natural products from these soil bacteria has steadily increased. To find novel classes of antibiotics, several researchers started to explore the chemistry of non-actinomycetes. In particular, myxobacteria and *Bacillus* spp. were found to produce several secondary metabolites with potential pharmaceutical applications [40-42]. Examples are the thuggacins [43] and myxovirescin [44], were isolated from myxobacteria. The lantibiotic mersacidin [45] has activity against methicillin-resistant *Staphylococcus aureus* and is produced by *Bacillus* species. Metagenomic analysis revealed that "as-yet-uncultured" microorganisms might be a

potential source of new antibiotics due to their huge biosynthetic potential [46]. Furthermore, approximately two-thirds of the bacterial phyla are still unexplored regarding their potential for the production of bioactive compounds to date. Among these neglected microbes are many predatory bacteria, which feed on other bacteria, including those that are resistant to known antibiotics [47]. Predatory behavior has been demonstrated to be associated with the production of lytic enzymes and bioactive compounds. Therefore, predatory bacteria might represent promising sources for natural product discovery. This hypothesis is supported by chemical analyses of predatory myxobacteria, which led to the identification of numerous antibacterial agents [48].

1.3 Predatory bacteria

Predatory bacteria are widely distributed in nature including rivers, soils, decaying woods, animal feces, ocean, sewage, and groundwater [49]. Occasionally, it has been observed that biosynthetic pathways for certain amino acid or vitamins are absent in predatory bacteria [50]. It is thus possible that predation has evolved to compensate for this deficiency. Predation plays an important role in regulating the composition of microbial communities. To date, predatory bacteria have been reported from seven phyla (Table 1).

Various predation strategies are known, which can be categorized into three groups: (i) epibiotic predation, (ii) endobiotic predation and (iii) group attack [51]. Predatory bacteria which exploit epibiotic predation attach to the prey surface and inject hydrolytic enzymes into prey cells. Representative bacteria pursuing this strategy are *Bdellovibrio exovorus, Micavibrio, Vampirovibrio* and *Nanoarchaeum* [50, 52-54]. Endobiotic predators penetrate the periplasmic space or the cytoplasm of the prey. *Bdellovibrio bacteriovorus* mainly utilizes this strategy [51]. Myxobacteria and *Herpetosiphon* attack in groups or so-called wolf packs. Here, a quorum of predatory cells is required to kill the prey organism by excreted hydrolytic enzymes and/or antibiotics [55]. Depending on the hydrolytic enzymes and antibiotics used, the predatory activity of wolf pack-forming bacteria can significantly vary, which is reflected in different degrees of swarm expansion on prey organisms.

Table 1. Taxonomic distribution of predatory bacteria and mode of predation.

Phylum	Class	Genera	Mode of predation
Actinobacteria	*Actinobacteria*	*Agromyces*[56]	Epibiotic
		**Streptomyces*[54, 57]	Epibiotic
		Streptoverticillium[58]	Epibiotic
Bacteroidetes	*Cytophagia*	*Cytophaga*[59]	Epibiotic
	Flavobacteriia	*Flavobacterium*[60]	Epibiotic
	Sphingobacteria	*Saprospira[61]*	Unknown
Chloroflexi	*Chloroflexia*	*Herpetosiphon*[62-63]	Wolf-pack
Cyanobacteria	*Melainabacteria*	*Vampirovibrio*[64]	Epibiotic
Firmicutes	*Bacilli*	*Paenibacillus*[65-67]	Epibiotic
Nanoarchaeota		*Nanoarchaeum*[49]	Epibiotic
Proteobacteria	*Alpha Proteobacteria*	*Ensifer*[68], *Micavibrio*[69]	Epibiotic
	Beta Proteobacteria	*Aristabacter*[49], *Cupriavidus*[70-71]	Epibiotic
	Gamma Proteobacteria	*Lysobacter*[72]	Wolf-pack
		Stenotrophomonas[73]	Epibiotic
	Delta Proteobacteria	*Bdellovibrio*[51-52]	Periplasmic
		Halobacteriovorax[74]	Endobiotic/Epibiotic
		Myxococcus and further genera belonging to the myxobacteria[64, 75]	Wolf-pack
		Peredibacter[76]	Unknown

* Not all bacteria in this genus are predatory bacteria

Among predatory bacteria, *Myxococcus xanthus* possesses a huge number of genes associated with hydrolytic enzymes such as proteases, phosphatases, lipases, and glucosidases [77-79]. A high number of hemolysin-related proteins are encoded in the genomes of *Micavibrio* that contribute to prey recognition and adhesion [53]. Gene expression analyses revealed that *Bdellovibrio* produces different hydrolytic enzymes during the predatory process [80]. Moreover, secondary metabolites with diverse biological activities have been proposed to play important roles in predation [48, 81]. For example, after the genetic inactivation of myxovirescin biosynthesis in *M. xanthus*, this bacterium was no longer capable to feed on *E. coli* [48].

Among all known predatory bacteria, actinobacteria and myxobacteria are already known to produce

secondary metabolites with pharmaceutical potential. Antibiotics from myxobacteria are considered to be relevant to predation activity [81]. Although *Bdellovibrio* and *Bdellovibrio*-like organisms (BALOS) are not antibiotics producers, they are able to prey on many human pathogens, such as *E. coli*, *Pseudomonas aeruginosa* and *Legionella* species [82-83]. Thus, they can be used as 'living antibiotics' in medicine and they are already used as biocontrol agents in agriculture, as well as cleaners for food equipment [84-85]. *B. bacteriovorus* has already been successfully used for of the treatment of chicken infected with *Salmonella* sp. [47].

1.4 The genus *Herpetosiphon*

The genus *Herpetosiphon* is the type genus of the family *Herpetosiphonaceae*, which belongs to the class *Chloroflexia* of the phylum *Chloroflexi* [86]. Unlike other bacteria in this class, the genus *Herpetosiphon* includes non-phototrophic chemoheterotrophic predatory bacteria. They exhibit similar morphological and phenotypic features as other *Chloroflexia.* They are multicellular filamentous bacteria with gliding motility [87]. The *Chloroflexia* are Gram-negative bacteria, although their cell wall shows several characteristics of Gram-positive bacteria. In particular, the peptidoglycan of *Herpetosiphon aurantiacus* features L-ornithine instead of diaminopimelic acid [88]. Currently, four species exist in the genus *Herpetosiphon*: *H. aurantiacus*, *H. geysericola*, *H. giganteus*, and *H. gulosus*. Since the latter two were described as part of this work, only *H. aurantiacus* and *H. geysericola* will be introduced here [63].

H. aurantiacus was first isolated in 1968 from lake water in USA, while *H. geysericola* originated from a hot spring in Mexico [62, 89-90]. *H. aurantiacus* is capable to utilize a variety of carbohydrates as sole carbon sources. *H. geysericola* has the distinctive cellulose hydrolysis ability [91]. A preliminary genome analysis of *H. aurantiacus* DSM 785 revealed several genes involved in secondary metabolite biosynthesis. A total of 14 biosynthetic loci were predicted from the genome of *H. aurantiacus,* of which only a single gene cluster could be associated with the production of a known compound class, i.e. the myxochelins [89]. The type species *H. geysericola* DSM 7119 contains nine putative

biosynthesis loci. Most of them show high similarity with biosynthetic gene clusters from *H. aurantiacus* [90]. In view of this, it can be expected that similar secondary metabolites are produced by these bacteria.

Although *H. aurantiacus* DSM 785 possesses a surprisingly large number of secondary metabolite biosynthetic gene clusters in relation to its genome size, only few secondary metabolites have been discovered from this species and the entire genus (Figure 1).

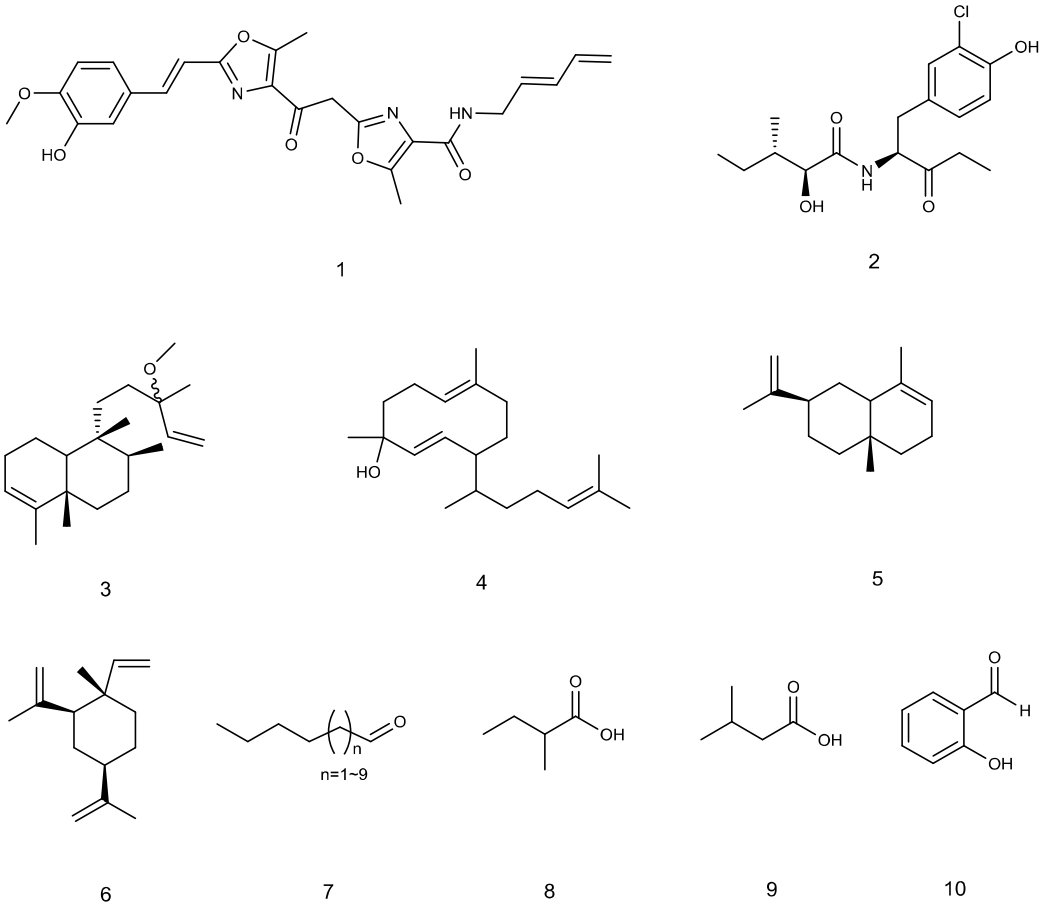


Figure 1. Natural products that have been isolated from *Herpetosiphon* spp.: 1, siphonazole [92-93]; 2, auriculamide [94]; 3, *O*-methylkolavelool [95]; 4, obscuronatin [96-97]; 5, α-selinene [97-98]; 6, β-elemene [97, 99]; 7, aldehydes; 8, 2-methylbutyric acid; 9, 3-methylbutyric acid; 10, salicyaldehyde [100].

Chemical analysis of the extracts from ten different *Herpetosiphon* strains resulted in the identification of the novel compound siphonazole [92]. Siphonazole and its methylated derivative

show cytotoxicity against human breast carcinoma (HTB-129) and human acute T-cell leukemia (TIB-152) [101]. Auriculamide is the first natural product from *H. aurantiacus* DSM 785. It was reported in 2015, and its biological function is not known [94]. The protein Haur_2145 encoded in the genome of *H. aurantiacus* DSM 785 shows high sequence similarity with the diterpene synthase Rv3377c from *Mycobacterium tuberculosis*. Heterologous expression led to the identification of *O*-methylkolavelool [95]. Similarly, heterologous expression of Haur_2987 in *Streptomyces avermitilis* resulted in the production of the diterpene obscuronatin. The Haur_2988 metabolic products β-elemene and α-selinene were also identified after heterologous expression in *S. avermitilis* [97]. In addition, some volatile compounds such as hexanal, tetradecanal, 3-methylbutyric acid, and salicylaldehyde were discovered from *H. aurantiacus* DSM 785 by headspace analysis [100].

1.5 Biosynthesis of natural products

Natural product biosynthesis involves comparatively few construction mechanisms. The structural diversity of natural products arises primarily from the combination of different building blocks. Most bacterial secondary metabolites are assembled by nonribosomal peptide synthetases (NRPSs), polyketide synthases (PKSs) and/or terpene synthases [102].

1.5.1 Polyketide biosynthesis

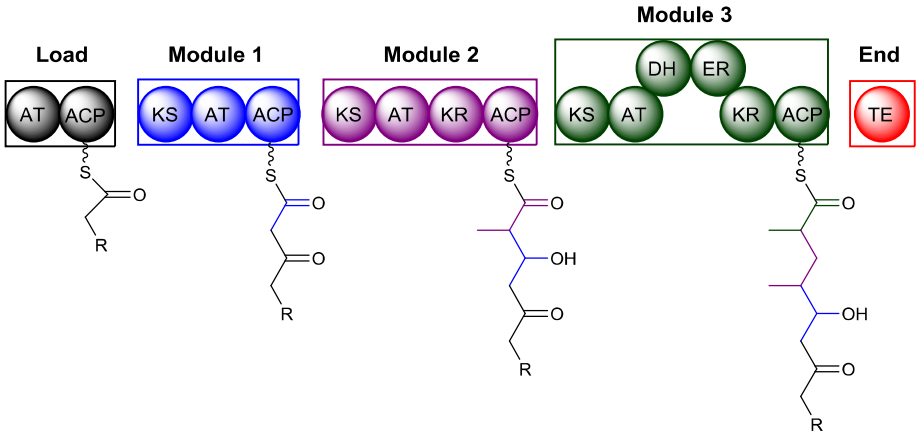
Polyketides are a family of natural products with a wide range of biological activities, such as antitumor, anticancer, cytotoxic and antioxidant. The enzymes that catalyze the biosynthesis of these compounds are divided into three distinct classes. Type I PKSs are modularly organized enzymes that are responsible for the biosynthesis of macrolides, polyethers and polyenes. A typical type I PKS module contains individual domains for selection, integration and modification of the acyl-CoA building blocks (Figure 2a). Acyltransferase (AT) domains are responsible for the selection of starter and extender units. They transfer the selected substrate to the phosphopantetheine arm of an acyl carrier protein (ACP). Ketosynthase (KS) domains catalyze a decarboxylative Claisen condensation

between an extender unit and the growing polyketide chain. Additional domains are involved in the reductive processing of the initially generated β-ketoacyl intermediates. A ketoreductase (KR) domain gives a secondary alcohol. A subsequent dehydration by a dehydratase (DH) domain yields an α,β-unsaturated acyl thioester, which can be further reduced by an enoylreductase (ER) domain. The terminal module of a type I PKS assembly line features a thioesterase (TE) domain, which catalyzes either hydrolysis or macrocyclization of the acyl chain [103-104].

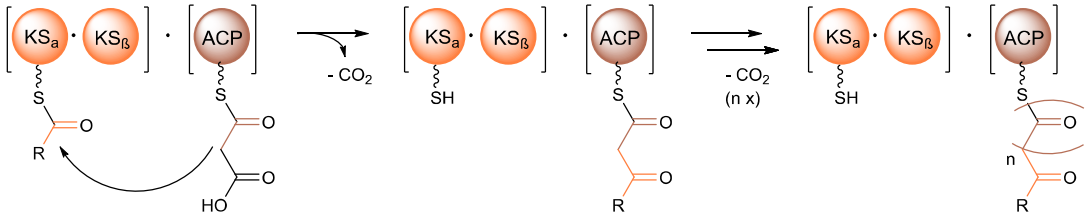
In general, type II PKSs are responsible for the formation of aromatic polyketides, such as tetracycline (antibacterial) [105], mensacarcin (antitumor) [106], pradimicins (antifungal) [107] and frenolicin B (antiparasitic) [108]. Each of these polycyclic aromatic compounds originates from the cyclization of a poly-β-ketoacyl chain. The biosynthesis of this linear precursor starts with an acyl-CoA unit (e.g., acetyl-CoA), which is subsequently transferred to the active site of a ketosynthase heterodimer (KS_α and KS_β) and then iteratively elongated via decarboxylative Claisen condensations with the extender unit malonyl-CoA (Figure 2b) [109].

Compared with type I and type II PKSs, type III PKSs are homodimeric iteratively acting synthases that only consist of a KS domain. Type III PKSs catalyze the biosynthesis of aromatic polyketides using a variety of acyl-CoA substrates as starter and extender units. Unlike type I and type II PKSs, the substrates of type III PKSs are not bound to ACP domains (Figure 2c) [110]. For example, 1,3,6,8-tetrahydroxynaphthalene (THN) is formed in *Streptomyces griseus* by the condensation of five malonyl-CoA substrates. Triketide pyrones of *Bacillus subtilis* are made by a type III PKS using long-chain fatty acyl-CoA thioesters as starter and malonyl-CoA as extender unit [111].

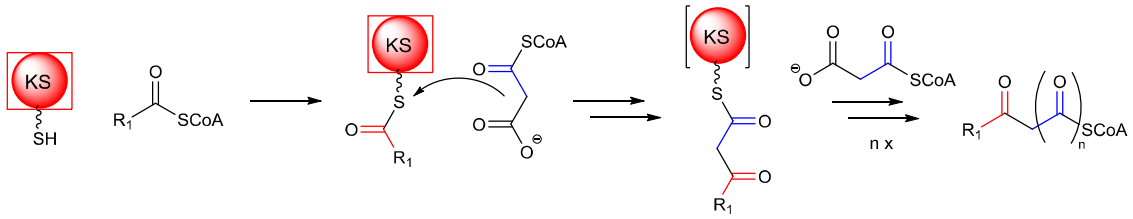
(a) Type I PKS



(b) Type II PKS



(c) Type III PKS



n=1, 2, 3, ...
R= C_nH_{2n+1}

Figure 2. Scheme of PKS catalytic reactions.

1.5.2 Non-ribosomal peptide biosynthesis

Nonribosomal peptide synthetases (NRPSs) are involved in the biosynthesis of antibacterial (daptomycin) [112], antifungal (bacillomycin) [113], iron-chelating (myxochelin) [114-115], immunosuppressive (cyclosporine A) [116] and antitumor (skyllamycin A) [117] compounds. Similar to type I PKSs, NRPSs possess a modular architecture of repeating catalytic domains. The latter typically incorporate amino acid building blocks into a growing peptide chain. The core catalytic functions are performed by condensation (C), adenylation (A) and peptidyl carrier protein (PCP) or thiolation (T) domains. The C domain is responsible for the peptide bond formation. It is located at the N-terminus of NRPS modules, except for the initiation module, which usually lacks this domain. The A domain catalyzes the formation of an aminoacyl thioester by activating and loading a selected amino acid onto the PCP domain. The PCP domain bears a phosphopantetheine (PPant) moiety, which has a key role in transferring the covalently bound aminoacyl and peptidyl substrates to the C domain. The PPant arm is installed post-translationally by a phosphopantetheinyl transferase. The final NRPS module of an NRPS assembly line typically contains a TE domain, which releases the nonribosomal peptide via hydrolysis or an intramolecular cyclization reaction [118-119].

In addition, each elongation module may contain additional domains that lead to structural diversity of NRPs. For example, epimerization (E) domains catalyze the conversion of L-amino acids into the corresponding D-amino acids. Cyclization (Cy) domains can replace C domains. They catalyze the incorporation and cyclodehydration of cysteine, serine or threonine units to give thiazoline, oxazoline or methyloxazoline rings. Vibriobactin [120] and yersiniabactin [121] are examples of natural products containing oxazoline and thiazoline residues, which were generated by Cy domains. These five-membered rings can be further oxidized to thiazoles or oxazoles by oxidation (Ox) domains. Methyltransferase (MT) domains mediate a methyl group transfer to carbon, nitrogen or oxygen atoms.

1.5.3 Isoprenoid biosynthesis

Terpenoids (also known as isoprenoids) represent the largest family of natural products that is known to date [122]. They possess diverse biological functions and may act, amongst others, as photosynthetic pigments (carotenoids), cell membrane constituents (cholesterol, ergosterol), or as antibiotics (e.g., pentalenolactone) [122-123].

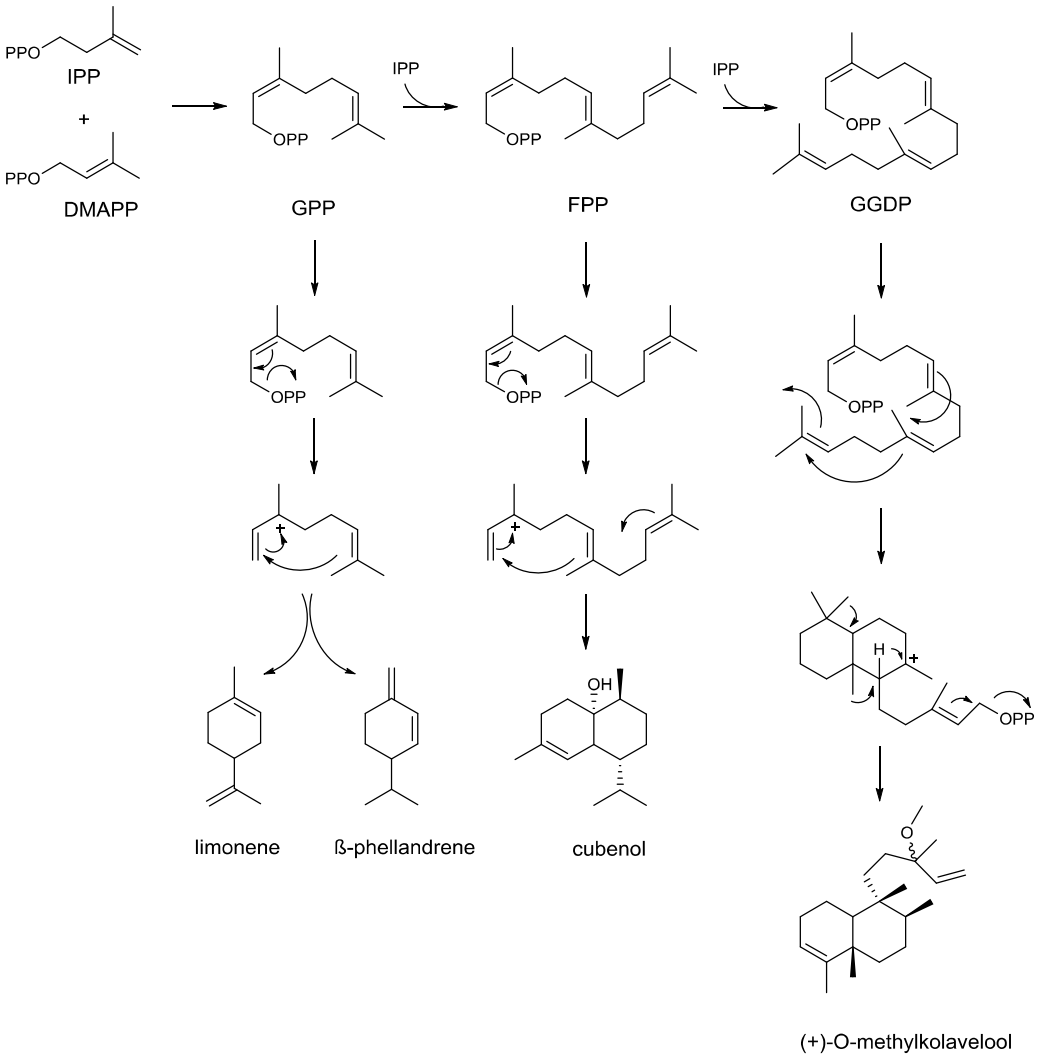


Figure 3. Examples of terpenoids generated from GPP, FPP and GGPP, respectively.

Although isoprenoids are structurally diverse, all of them derive from the common five-carbon (C_5) building blocks isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). These building blocks can be generated by two independent pathways: the mevalonate (MEV) pathway or, alternatively, the 2C-methyl-D-erythritol 4-phosphate (MEP) pathway [123]. The different isoprenoid backbones arise from the head-to-tail coupling of the isoprene building blocks. Monoterpenes derive from geranyl diphosphate (GPP, C_{10}), which is assembled from DMAPP and IPP. Another condensation of GPP with IPP unit results in the production of farnesyl diphosphate (FPP, C_{15}), which provides the basic carbon skeleton of all sesquiterpenes. FPP can also be further converted to the diterpene precursor geranylgeranyl diphosphate (GGPP, C_{20}) by head-to-tail condensation with IPP. This process can be further continued to the sesterterpene precursor geranylfarnesyl diphosphate (GFPP, C_{25}) and so on [124-125]. Further structural diversity is generated through cyclization reactions, which are carried out by terpene cyclases [126] (Figure 3).

1.6 Genome Mining

Bioactivity-guided fractionation, which has been traditionally used for drug discovery, is not only a time-consuming method, but also increasingly leads to the reisolation of known natural products. This problem is quite obvious in the case of actinomycetes, which have been chemically investigated over a long period [127]. In contrast, the computational mining of genomes offers improved chances for the discovery of new natural products [128]. Genome mining refers to multimethodology, including homology sequence identification, bioinformatic prediction of gene and pathway function, gene expression and biochemical analysis, as well as the targeted isolation of new metabolites [128-131]. Various methods and tools have been developed to facilitate genome mining studies. To identify putative biosynthetic gene clusters, a number of computational sequence comparison tools are available, such as Bagel3 [132], antiSMASH [133], and RiPPMiner [134]. The model actinomycete *Streptomyces coelicolor* A3 was one of the first bacteria, which was subjected to extensive genome mining studies [135-136]. Recently, genome mining has been used on a large scale for searching polyketides, lanthipeptides, isoprenoids, and siderophores biosynthesis genes in more than 800

actinobacterial genomes [137].

Since several biosynthesis genes remain silent under laboratory cultivation conditions, the number of known secondary metabolites is significantly lower than that predicted by genome mining [138-139]. Activation of silent genes for the identification of novel secondary metabolites has thus attracted major interest. Gene overexpression is frequently used to activate, produce, and modify the pathways of natural products [140]. This strategy can involve the activation of silent gene clusters by genetic manipulation of the producing organism, or the cloning of entire biosynthetic clusters in a suitable heterologous host. However, cloning and efficient expression of the corresponding biosynthetic gene cluster still remain challenging due to the large size of many biosynthetic gene clusters or the ineffectiveness of host cells [141].

Alternatives to pathway overexpression are the one strain-many compounds (OSMAC) and the genomisotopic approach (Figure 4). In the OSMAC approach the accessible number of secondary metabolites from a given microbial source is increased by systematically altering cultivation parameters, such as media composition, temperature, and aeration. The complex bacterial metabolic network is known to respond to different cultivation conditions. This offers opportunities for inducing changes in metabolite profiles [142-143]. For example, the use of different media supplements in *Streptomyces* sp. A1 led to the production of streptazoline and its analogous [143]. Two antibacterial compounds, an eudesmane sesquiterpenoid and butanolide E, were isolated from *Lentzea violacea* AS08 by changing the composition of the growth media [144].

The genomisotopic approach aims at the recovery of peptidic natural products using an isotope-guided fractionation by LC-MS or NMR. An appropriate isotopically labeled precursor is selected for a feeding experiment with the aid of genome sequence analysis. After its incorporation into the target compound, the specific isotope labeling pattern can be used to trace and to isolate the natural product. This approach has been used successfully to isolate cyclic lipopeptides from *Pseudomonas fluorescens*. Following genome sequence analysis, an orphan gene cluster was identified. Feeding the optimal isotopically labeled precursor based on this analysis resulted in the discovery of the orfamides [145].

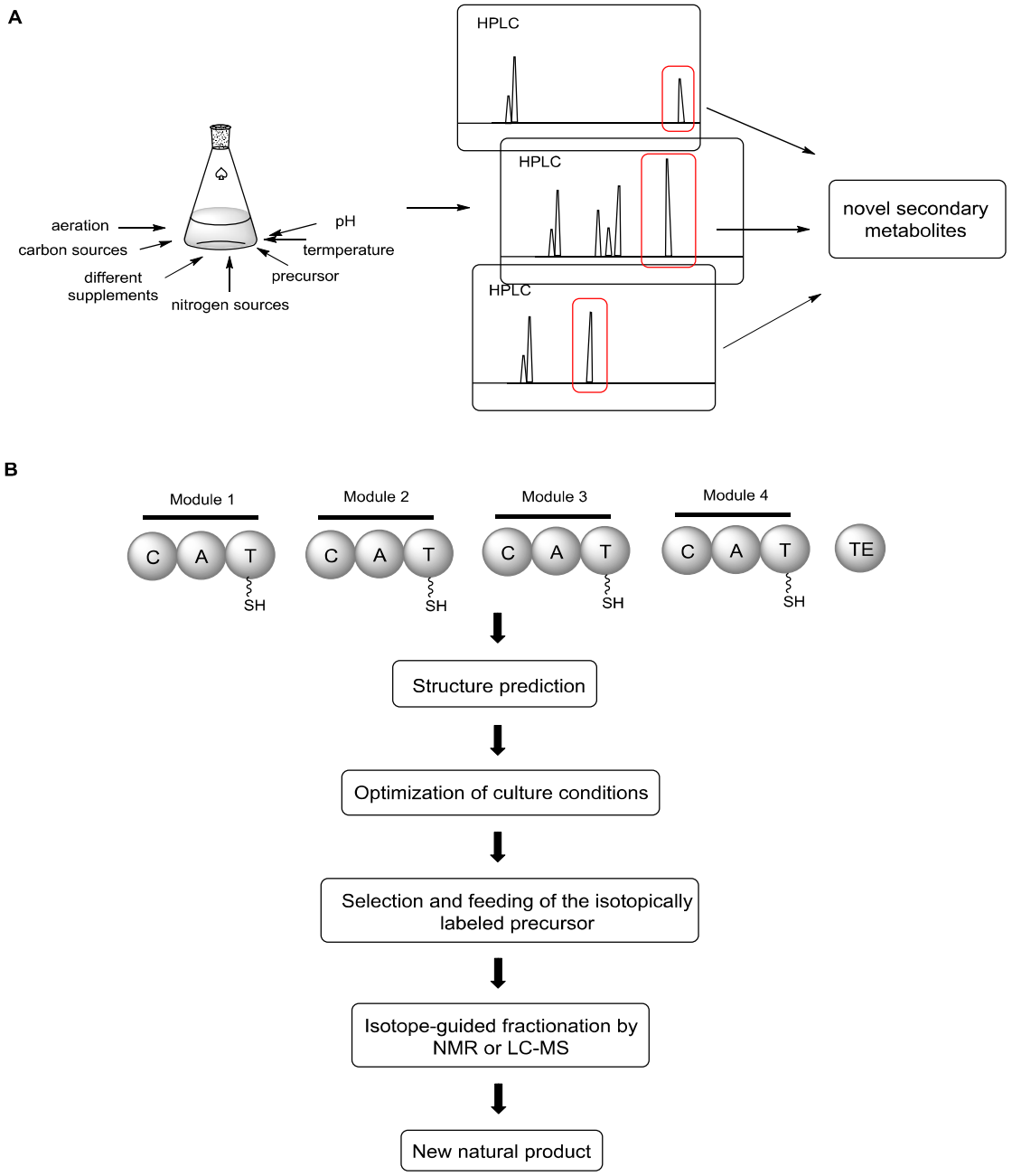


Figure 4. General concept of the OSMAC (A) and of the genomisotopic (B) approach.

2 Scope of the project

The genus *Herpetosiphon* is represented by bacteria, which feed on other microbes, and are further capable of natural product biosynthesis. Taxonomically, these predatory bacteria belong to the phylum *Chloroflexi*, which is known to harbor metabolically and physiologically highly diverse organisms. Thus, a genomic comparison of *Herpetosiphon* and other *Chloroflexi* species could provide information on the distribution of biosynthetic pathways in the entire phylum. Furthermore, it would be possible to test whether there is a possible correlation between the presence of biosynthetic genes and the predatory lifestyle of *Herpetosiphon* spp. More than hundred *Herpetosiphon*-like strains are stored in different strain collections all over the world. However, only few of these organisms have actually been subject to a thorough characterization concerning their phylogenetic position or predatory activity. The scope of this PhD project was thus:

(1) To identify natural products from *Herpetosiphon* spp. and to investigate their antimicrobial properties,

(2) To clarify the taxonomic standing and physiological properties of *Herpetosiphon*-like strains, and

(3) To uncover commonalities and differences in the biosynthetic potential of *Herpetosiphon* and other *Chloroflexi* bacteria.

3 Publications

Manuscript A:

Xinli Pan, Nicole Domin, Sebastian Schieferdecker, Hirokazu Kage, Martin Roth, and Markus Nett

Herpetopanone, a diterpene from *Herpetosiphon aurantiacus* discovered by isotope labeling. *Beilstein J. Org. Chem.* **2017,** *13,* 2458–2465

The genome of *Herpetosiphon aurantiacus* DSM 785 harbors numerous terpene pathways. To identify their products, the predatory bacterium was grown in the presence of ^{13}C-labeled glucose. The metabolism of the latter induced characteristic mass shifts in the terpenome of *H. aurantiacus*, which were used to isolate a previously unknown diterpene.

Xinli Pan, Nicole Domin and Sebastian Schieferdecker performed laboratory-scale cultivations of *H. aurantiacus* DSM 785, as well as the isolation and chemical analysis of herpetopanone. Furthermore, X. Pan conducted and analyzed the results of the feeding study. Dr. Hirokazu Kage investigated the function of the terpene synthase genes in heterologous expression studies. Dr. Martin Roth carried out a batch fermentation of *H. aurantiacus* DSM 785 with ^{13}C-labeled glucose. Prof. Dr. Markus Nett designed the research project, elucidated the structure of herpetopanone and wrote the final manuscript. Own contribution (Xinli Pan): 50%.

Manuscript B:

Xinli Pan, Hirokazu Kage, Karin Martin, and Markus Nett

*Herpetosiphon gulosu*s sp. nov., a filamentous predatory bacterium isolated from sandy soil and *Herpetosiphon giganteus* sp. nov., nom. rev. *Int. J. Syst. Evol. Microbiol.* **2017**, 67, 2476-2481

Genotypic and phenotypic analyses of *Herpetosiphon* strains, which were obtained from the German Collection of Microorganisms and Cell Cultures, led to the description of a new species in this genus, *H. gulosus*. Furthermore, the species name *H. giganteus* was revived. Xinli Pan performed the cultivation of *H. gulosus* DSM 52871 and *H. giganteus* DSM 589, as well as 16S rRNA gene sequence,

fatty acid, morphology and predation analyses. Karin Martin performed the menaquinone composition analyses, provided material for antibiotics and chemotaxonomic analyses. Dr. Hirokazu Kage provided support in the interpretation of sequencing results. Prof. Dr. Markus Nett designed the research project and wrote the final manuscript. Own contribution (Xinli Pan): 90%.

Manuscript C:

Xinli Pan and Markus Nett

Natural product biosynthesis in the phylum *Chloroflexi*: a genomics perspective.

(Manuscript in preparation)

The genetic potential for secondary metabolites production was investigated in the phylum *Chloroflexi*. Species from the genera *Herpetosiphon*, *Ktedonobacter* and *Thermogemmatispora* have the ability to produce known or unknown secondary metabolites.

Xinli Pan and Prof. Dr. Markus Nett performed bioinformatic analyses. Xinli Pan wrote the final manuscript. Own contribution (Xinli Pan): 95%.

Prof. Dr. Markus Nett

3.1 Manuscript A

Xinli Pan, Nicole Domin, Sebastian Schieferdecker, Hirokazu Kage, Martin Roth, and Markus Nett

Herpetopanone, a diterpene from *Herpetosiphon aurantiacus* discovered by isotope labeling. *Beilstein J. Org. Chem.* **2017,** *13,* 2458–2465

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Herpetopanone, a diterpene from *Herpetosiphon aurantiacus* discovered by isotope labeling

Xinli Pan[1,2], Nicole Domin[2], Sebastian Schieferdecker[2], Hirokazu Kage[1], Martin Roth[2] and Markus Nett[*1]

Full Research Paper

Open Access

Address:
[1]Department of Biochemical and Chemical Engineering, Technical Biology, Technical University Dortmund, Emil-Figge-Strasse 66, 44227 Dortmund, Germany and [2]Leibniz Institute for Natural Product Research and Infection Biology, Hans Knöll Institute, Beutenbergstr. 11a, 07745 Jena, Germany

Email:
Markus Nett[*] - markus.nett@bci.tu-dortmund.de

* Corresponding author

Keywords:
genome mining; herpetopanone; *Herpetosiphon*; isotope labeling; terpene

Beilstein J. Org. Chem. **2017,** *13,* 2458–2465.
doi:10.3762/bjoc.13.242

Received: 15 August 2017
Accepted: 30 October 2017
Published: 17 November 2017

Associate Editor: C. Stephenson

Abstract

The genome of the predatory bacterium *Herpetosiphon aurantiacus* 114-95^T harbors a number of biosynthesis genes, including four terpene cyclase genes. To identify the terpenes biosynthesized from *H. aurantiacus* 114-95^T, we fed the strain with ^{13}C-labeled glucose and, subsequently, searched for characteristic mass shifts in its metabolome. This approach led to the discovery of a new natural product, of which the isotope pattern is indicative for a diterpene originating from the methylerythritol phosphate pathway. After large-scale fermentation of *H. aurantiacus* 114-95^T, the putative diterpene was isolated in sufficient quantity to enable NMR-based structure elucidation. The compound, for which the name herpetopanone is proposed, features a rare octahydro-1*H*-indenyl skeleton. Herpetopanone bears resemblance to cadinane-type sesquiterpenes from plants, but is structurally entirely unprecedented in bacteria. Based on its molecular architecture, a possible biosynthetic pathway is postulated.

Introduction

Terpenoids represent the largest group of natural products with about 60,000 different compounds being known. They occur in all three domains of life and are known to fulfill a variety of different functions, e.g., as membrane constituents, chemical attractants or feeding deterrents [1]. Over a long period, terpenoids were mainly reported from plants and, to a lesser degree, also from fungi and marine invertebrates. In recent years, however, the discovery of terpenoids from prokaryotes has gained momentum. The ease of DNA sequencing has strongly favored this development, contributing to the identification of numerous bacterial terpene cyclase genes [2,3]. Both in vitro and in vivo approaches involving recombinant enzymes are commonly pursued for their functional characterization [4]. Care must be taken, however, in interpreting the results of these

analyses, as the products of terpene cyclases are often subject to enzymatic modifications in their native producers. The exclusive testing of a terpene cyclase might hence unveil a biosynthetic intermediate rather than the final product of a secondary metabolite pathway [5,6]. To avoid this problem, we here describe a method for the identification of terpenoids in their natural bacterial hosts, which is based on the feeding of isotopically labeled glucose.

The linear oligoprenyl units, which constitute the carbon backbones of terpenoids, arise from the condensation of activated isoprene units, namely isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). The latter two precursors are synthesized by either the mevalonate (MEV) or methylerythritol phosphate (MEP) pathway [7]. Both the MEV and MEP pathway branch from glycolysis. Depending on the respective route, the metabolism of singly labeled glucose gives rise to a characteristic carbon labeling pattern in IPP and DMAPP (Figure 1) [8]. This feature has proven extremely useful to unravel complex cyclization cascades and carbon–carbon rearrangements in the biosynthesis of some terpenoids [9-12]. We anticipated that the resulting mass shifts could also be valuable in the field of natural product discovery. By comparing the ion chromatograms of cultures that were grown in the presence or absence of singly labeled glucose, it might be possible to identify terpenoid natural products in a complex metabolic background.

To validate the feasibility of this approach, we chose the predatory bacterium *Herpetosiphon aurantiacus* 114-95^T as a test organism. This strain is capable to produce a variety of polyketides and nonribosomal peptides [13-15] and possesses pathways to supply specific building blocks for these natural products [16]. Furthermore, genomic analyses revealed that *H. aurantiacus* 114-95^T features four genes coding for putative terpene cyclases, i.e., Haur_2145, Haur_2987, Haur_2988, and Haur_4149. While the class II cyclase Haur_2145 had already been associated with the production of the terpenoid *O*-methylkolavelool [17], the products of the other three enzymes have not been identified from *H. aurantiacus* 114-95^T.

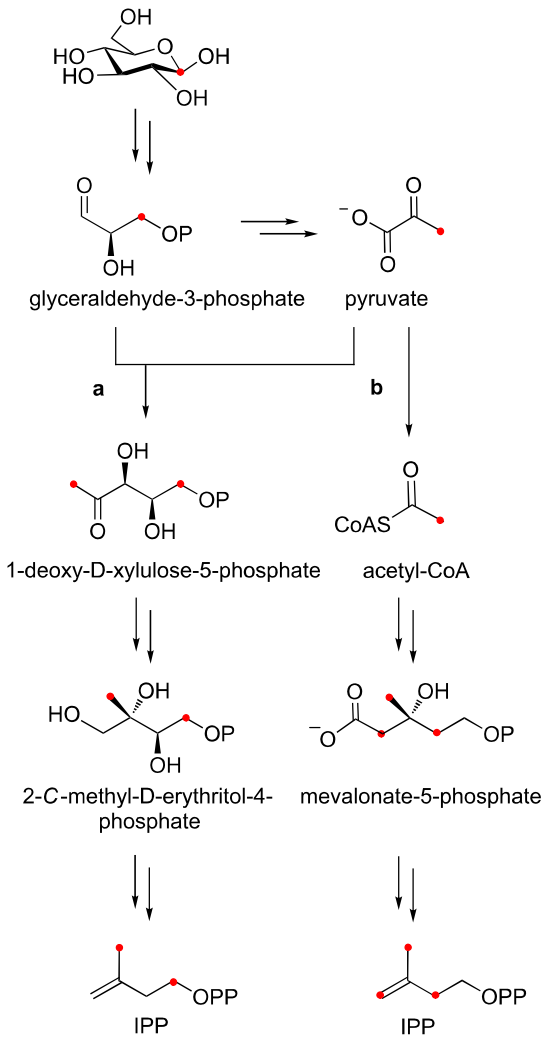


Figure 1: Distribution of isotopic labels from [1-^{13}C]-glucose via the MEP (route a) and MEV pathway (route b). Potentially labeled carbon atoms are highlighted with red dots.

Results and Discussion

For the metabolic labeling experiment, *H. aurantiacus* 114-95^T was grown in modified Van Niel's yeast (VNY) medium. We had previously observed that *Herpetosiphon* cultures grown in this medium develop a deep orange color due to an increased production of carotenoids and it was hence speculated that VNY medium might also support the biosynthesis of other terpenoids. The cultivation was conducted for 7 days in the presence of [1-^{13}C]-labeled D-glucose. Cultures supplemented with non-labeled D-glucose served as a control. Extracts from the different *Herpetosiphon* cultures (labeled vs control) were subjected to LCMS analyses. From the comparison of the respective spectra, a possible terpene was identified. The candidate compound lacked absorption maxima in the visible region of the spectrum and could therefore not represent a carotenoid. In our control cultures, the compound exhibited a pseudomolecular ion at *m/z* 341.2 $[M + H]^+$. After feeding of [1-^{13}C]-glucose, however, the same molecule showed a complex isotope pattern with a gradual increase of its *m/z* value up to a maximum of 8 Da (Figure 2).

Two conclusions were drawn from this observation. First, it was evident that the isotope labeling was incomplete and that the incorporation of non-labeled metabolic intermediates had led to a mixture of isotopologues and isotopomers. Second, the maximal incorporation of eight labeled carbons was not possible if the

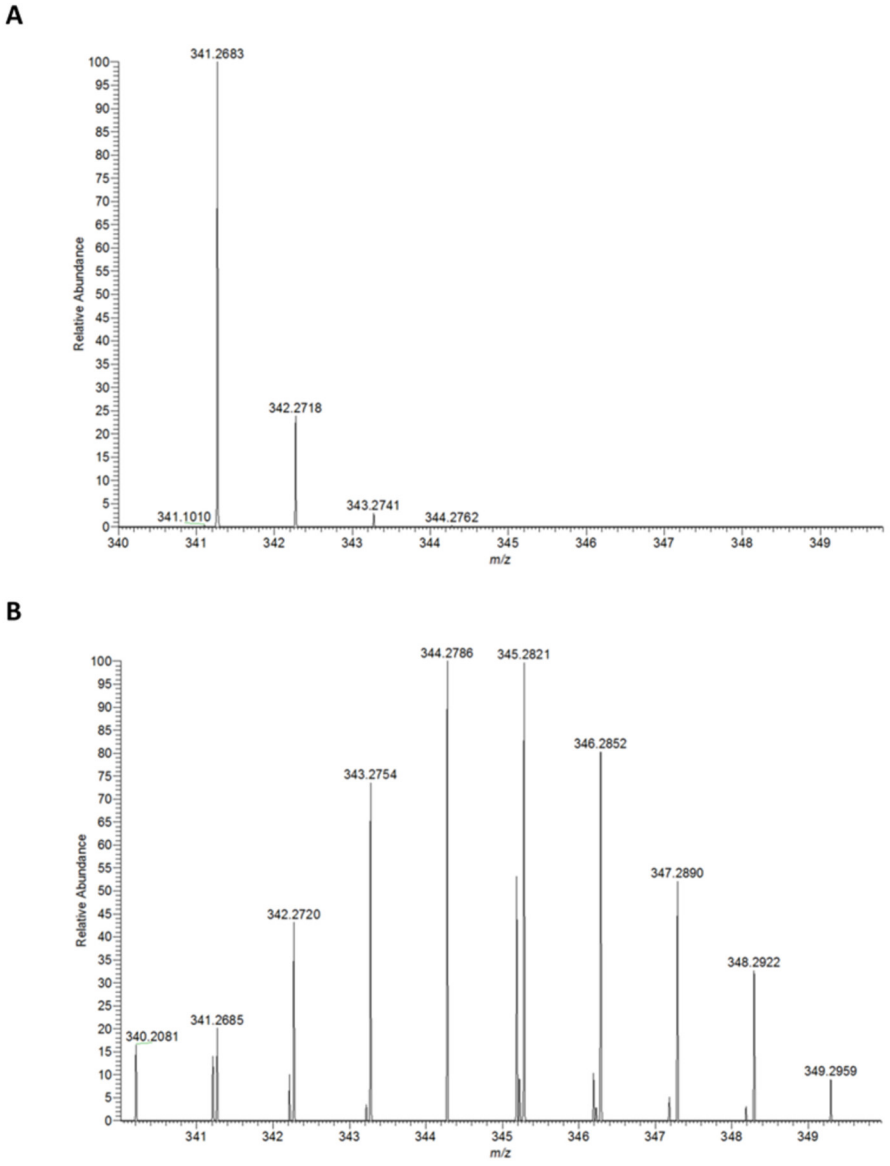


Figure 2: High-resolution mass spectra of a metabolite from *H. aurantiacus* obtained after feeding of unlabeled (A) and [1-^{13}C]-labeled D-glucose (B).

compound originated from the MEV pathway (Figure 1). The possible number of labeled carbon atoms was compatible, however, with an octaketide origin or with a diterpene from the MEP pathway. We then recorded the high resolution (HR) mass of the unlabeled compound and determined its molecular formula as $C_{20}H_{36}O_4$. Although the elemental composition does not necessarily exclude a polyketide origin, it perfectly matches a diterpene comprising four intact isoprene units.

To obtain sufficient material for structure elucidation, the fermentation of *H. aurantiacus* 114-95^T was repeated on a 50 L scale in VNY medium supplemented with non-labeled D-glucose. The resulting culture broth was centrifuged, and the supernatant was treated with the resin XAD-2 to adsorb the metabolites that had been secreted during the cultivation. After extraction of the adsorber resin with methanol, a preliminary fractionation was accomplished by flash column chromatography on RP silica gel. Fractions containing the target compound were identified by LC–MS, pooled and further purified via semipreparative RP–HPLC to give 19 mg of herpetopanone (**1**).

The IR spectrum of **1** revealed two distinctive bands at 3395 and 1699 cm^{-1}, which were assigned to hydroxyl and carbonyl stretching vibrations, respectively. The molecular formula of **1**, which had already been established as $C_{20}H_{36}O_4$, corresponds to three degrees of unsaturation. The signals in the ^{13}C NMR

spectrum were allocated by DEPT measurements to five methyl, six methylene and six methine groups, as well as three quaternary carbons (Table 1). Only a single carbon (C-19) was sp^2-hybridized; its resonance at δ 212.2 ppm indicated the presence of a ketone group. It was hence concluded that the structure of **1** must comprise two ring structures in order to comply with the required degrees of unsaturation. Consolidating IR, MS and DEPT data, three hydrogen atoms were assigned to hydroxy groups. The corresponding groups were placed next to C-2, C-4 and C-12 on the basis of ^{13}C chemical shifts.

While the methyl protons of **1** were readily identified in the ^1H NMR spectrum, the assignment of several other protons and the readout of their coupling constants was challenging due to severe signal overlapping. A TOCSY spectrum suggested that, with the exception of four methyl groups (CH_3-1, CH_3-3, CH_3-13, CH_3-20), all protons were part of the same spin system. Structure elucidation started with CH_2-16 and CH-17, because the protons of these two adjacent groups showed discrete signals in the ^1H NMR spectrum. Their multiplicities as well as the associated *J* values indicated further vicinal coupling partners. In case of H-16a and H-16b, correlations were also observed with H-15a and H-15b in the COSY spectrum, while H-18 was identified as a coupling partner of H-17. Homonuclear correlations from H-18 to H-9 and H-14 enabled us to expand the deduced partial structure, which was then verified through HMBC and HSQC data. Heteronuclear long-range interactions established the missing substituent at C-17 as an acetyl group and allowed the closure of the cyclopentyl ring in **1** (Figure 3). The proton signal of CH_3-13 appeared as a singlet, thereby excluding a hydrogen-bearing carbon as an immediate neighbor. In the HMBC spectrum, H_3-13 showed three heteronuclear correlations, of which only one occurred with a quaternary carbon (C-12). It was hence evident that the methyl group is attached to the hydroxy-substituted C-12, which itself is further connected with C-11 and C-14. HMBC interactions from H-15a and H-15b to C-12 lend additional support for the placement of the latter. The protons of the methylene group in position 11 are magnetically non-equivalent. This feature led to the pairwise occurrence of correlation peaks in the HMBC spectrum and promoted the identification of their ^{13}C coupling partners. In addition to the expected interactions with C-13 and C-14, H-11a and H-11b showed correlations to C-9 and C-10. From the observation that H-18 exhibited an HMBC correla-

Table 1: NMR spectroscopic data of herpetopanone (**1**) in chloroform-d_1.[a]

position	δ_C [ppm]	δ_H [ppm], M (*J* in Hz)	HMBC
1	23.1, CH_3	1.14, s	2, 3, 4
2	73.7, C_q		
3	26.4, CH_3	1.19, s	1, 2, 4
4	79.0, CH	3.31, dd (9.8, 2.1)	1, 2, 3, 5, 6
5	29.8, CH_2	a: 1.43, nr b: 1.19, nr	
6	33.1, CH_2	a: 1.46, nr b: 1.20, nr	5, 8 5, 8
7	34.9, CH	1.19, nr	8
8	13.9, CH_3	0.68, d (6.2)	6, 7, 9
9	47.6, CH	1.14, nr	8
10	23.2, CH_2	a: 1.52, nr b: 1.10, nr	9, 11 9, 11
11	41.7, CH_2	a: 1.79, nr b. 1.36, nr	9, 13, 14 9, 10, 13, 14
12	73.8, C		
13	20.0, CH_3	1.18, s	11, 12, 14
14	56.7, CH	1.49, nr	13
15	25.2, CH_2	a: 1.80, nr b: 1.40, nr	12, 14, 18 12, 14
16	28.5, CH_2	a: 1.94, dddd (13.3, 11.7, 10.0, 8.0) b: 1.56, dddd (13.3, 9.3, 5.5, 1.5)	14, 15, 17, 18, 19 14, 15, 17, 18, 19
17	55.9, CH	2.59, ddd (11.7, 9.5, 5.6)	9, 15, 16, 18, 19, 20
18	46.4, CH	1.81, nr	7, 10, 16
19	212.2, C_q		
20	29.3, CH_3	2.16, s	16, 17, 18, 19

[a]nr, not resolved (chemical shift was deduced from the HSQC spectrum).

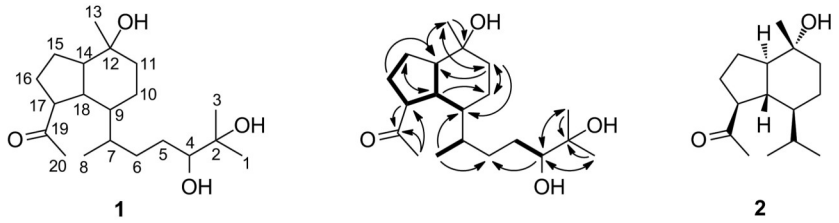


Figure 3: Structures of herpetopanone (**1**) and oplopanone (**2**), as well as selected COSY (bold lines) and HMBC (arrows) interactions in **1**.

tion to C-10, but not to C-11, we deduced the octahydro-1*H*-indene scaffold of **1**. The highly shielded protons of CH_3-8 exhibited HMBC correlations to C-6, C-7, and C-9. Although their vicinal coupling partner could not be unequivocally assigned in the COSY spectrum, the splitting of the corresponding signal was only consistent with a methine being attached to a methyl group. Since CH_3-8 could not be placed next to CH-9, it had to be linked with CH-7. The two methyl groups, CH_3-1 and CH_3-3, appeared as singlets in the ^{1}H NMR spectrum. They were positioned next to the quaternary carbon C-2 on the basis of HMBC interactions. Interpretation of 2D NMR data allowed an extension of this residue with CH-4 and CH_2-5. Ultimately, a heteronuclear long-range correlation from H-4 to C-6 linked the two fragments to give the two-dimensional structure of **1**. Overlapping ^{1}H signals led to several assignment ambiguities in the NOESY spectrum and, thereby, impeded the deduction of the relative configuration. Attempts to crystallize **1** are currently performed in our laboratory.

Following its isolation, **1** was evaluated in an agar diffusion assay against a standard selection of Gram-positive and Gram-negative bacteria as well as some fungi. However, the compound was inactive at all tested concentrations.

A literature search revealed that **1** is the first microbial terpene possessing an octahydro-1*H*-indene backbone, whereas a plant-derived sesquiterpene with this feature was already reported in 1965. Oplopanone (**2**, Figure 3) was originally isolated from *Oplopanax japonicus* [18,19], but can also be found in a number of other plants as well as red algae [20,21]. Comparison of the chemical shifts in **1** with published NMR data for **2** [22] confirmed our assignment of the octahydro-1*H*-indene skeleton.

The biosynthesis of the bicyclic ring system in **2** was previously proposed to occur via an α-cadinol intermediate, which undergoes a ring contraction reaction [19]. The cadinane family of sesquiterpenes, which also includes α-cadinol, originates from germacrene D [7,23]. In the case of **1**, an analogous pathway can be postulated, which is depicted in Figure 4. The biosynthesis would hence start with geranylgeranyl pyrophosphate (GGPP). Upon ionization, the double bond nearest the diphosphate can adopt a *Z* configuration, thereby facilitating an intramolecular cyclization to a cyclodeca-1,5-diene by electrophilic attack of the allylic carbocation onto the corresponding double bond. A 1,3-hydride shift by Wagner–Meerwein rearrangement followed by another cyclization would then give rise to an octahydronaphthalene cation. Eventually, the addition of OH^- would lead to an α-cadinol-type diterpene. For the ring contraction, we would propose a two-step sequence of epoxidation and rearrangement [24], which would directly lead to the acetyl group in **1**. Recently, heterologous expression of the terpene cyclase Haur_2987 in an actinomycete led to a product, which was identified as the soft coral-derived diterpene obscuronatin [25,26]. The biosynthesis of this compound can be easily rationalized via the proposed herpetopanone pathway (route b in Figure 4). Following the formation of the cyclodeca-1,5-diene intermediate, two successive 1,3-hydride shifts and a final addition of OH^- would yield obscuronatin (**3**). Interestingly, **3** was observed to undergo facile water elimination to give compound **4** [26]. Obscuronatin might thus serve as a biosynthetic intermediate or, alternatively, as a shunt product in the herpetopanone pathway. An involvement of Haur_2987 in the biosynthesis of **1** seems highly likely.

Conclusion

In summary, we have identified a new diterpene from the predatory bacterium *H. aurantiacus* 114-95^T using an isotope labeling strategy and resolved its structure. A plausible pathway for the biosynthesis of herpetopanone was deduced from its molecular architecture. Furthermore, candidate genes for the necessary ring formations were identified and will be the subject of future investigations.

Experimental

General experimental procedures

The IR spectrum was recorded on a benchtop FT-IR 4100 spectrometer (JASCO). HRMS analyses were carried out using an Exactive Mass Spectrometer (Thermo-Scientific). NMR spectra were measured at 300 K on Bruker Avance III spectrome-

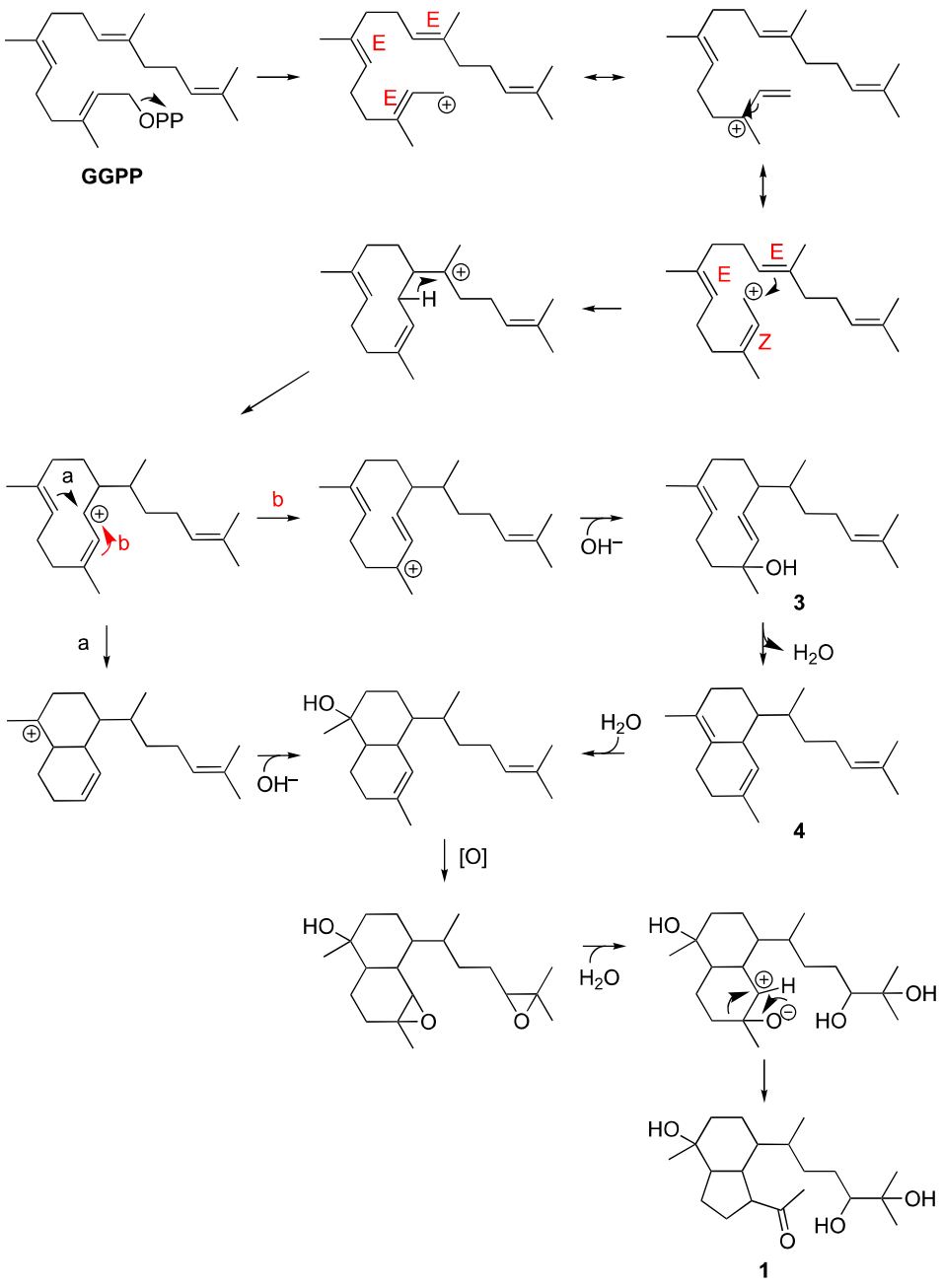


Figure 4: Proposed biosynthesis of **1** via two alternative routes (a) and (b). Route (b) involves the known diterpene obscuronatin (**3**).

ters with chloroform-d_1 (δ_H 7.24 ppm; δ_C 77.0 ppm) as solvent and internal standard.

Strain and growth conditions

H. aurantiacus 114-95^T was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ). The bacterium was routinely cultured at 30 °C in CY (casitone 3 g/L, yeast extract 1 g/L, $CaCl_2{\cdot}2\ H_2O$ 1 g/L, pH 6.8) or VNY liquid medium (sodium glutamate 10 g/L, yeast extract 2 g/L, $MgSO_4{\cdot}7\ H_2O$ 2 g/L, NaH_2PO_4 0.10 g/L, Na_2HPO_4 0.05 g/L, pH 6.5) under oxic conditions with gentle shaking (90 rpm).

Feeding experiments

Feeding experiments were conducted in 2 L Erlenmeyer flasks containing 1 L of VNY liquid medium. The precursors, [1-^{13}C]-D-glucose and non-labeled D-glucose, were added aseptically

as filter-sterilized aqueous solutions (0.05 M) at the time of inoculation. The fermentation was conducted at 30 °C for 7 days. After centrifugation and removal of the biomass, the supernatant was treated with 10% (w/v) XAD-2 resin (Sigma-Aldrich) in order to adsorb the secreted metabolites. The resin was thoroughly washed with distilled water and, subsequently, eluted with three bed volumes of methanol. The eluates were concentrated under vacuum and the resulting extracts were subjected to LCMS analysis.

Production and isolation of herpetopanone

The fermentation and extraction conditions were the same as described in the feeding experiments, except that *H. aurantiacus* 114-95^T was grown on a 20 L scale in VNY medium and that the cultures were only supplemented with non-labeled D-glucose. The extract that was obtained after the XAD-2 elution was dissolved in 20% (v/v) aqueous MeOH and fractionated by flash column chromatography over Polygoprep C18 (Macherey-Nagel) using an increasing concentration of MeOH in water. Fractions containing **1** were identified by LC–MS analysis, pooled and purified by reversed-phase HPLC on a Shimadzu UFLC liquid chromatography system equipped with a Nucleodur C18 HTec column (VP 250 × 10 mm, 5 µm; Macherey-Nagel) using a gradient from 10% (v/v) MeOH in H_2O (+ 0.1% trifluoroacetic acid) to 100% (v/v) MeOH over 30 min. The purity of the isolated herpetopanone determined by HPLC was >95% at all wavelengths (compared to sum of total peak areas).

Agar diffusion assay

The assay was performed as previously described [27]. The test organisms included *Bacillus subtilis* ATCC 6633, *Staphylococcus aureus* SG 511, *Mycobacterium vaccae* IMET 10670, *Escherichia coli* SG 458, *Pseudomonas aeruginosa* K 799/61, *Sporobolomyces salmonicolor* SBUG 549, *Candida albicans* ATCC 14053 and *Penicillium notatum* JP 36.

Supporting Information

Supporting Information File 1

IR and NMR spectra of herpetopanone. Metabolic profile of *H. aurantiacus* 114-95^T.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-13-242-S1.pdf]

Acknowledgements

X. Pan thanks the China Scholarship Council (CSC) for a doctoral stipend. We gratefully acknowledge H. Heinecke (Leibniz Institute for Natural Product Research and Infection Biology) for recording the IR spectrum of herpetopanone.

ORCID® iDs

Markus Nett - https://orcid.org/0000-0003-0847-086X

References

1. Köksal, M.; Hu, H.; Coates, R. M.; Peters, R. J.; Christianson, D. W. *Nat. Chem. Biol.* **2011,** *7,* 431–433. doi:10.1038/nchembio.578
2. Cane, D. E.; Ikeda, H. *Acc. Chem. Res.* **2012,** *45,* 463–472. doi:10.1021/ar200198d
3. Daum, M.; Herrmann, S.; Wilkinson, B.; Bechthold, A. *Curr. Opin. Chem. Biol.* **2009,** *13,* 180–188. doi:10.1016/j.cbpa.2009.02.029
4. Dickschat, J. S. *Nat. Prod. Rep.* **2016,** *33,* 87–110. doi:10.1039/C5NP00102A
5. Zhao, B.; Lin, X.; Lei, L.; Lamb, D. C.; Kelly, S. L.; Waterman, M. R.; Cane, D. E. *J. Biol. Chem.* **2008,** *283,* 8183–8189. doi:10.1074/jbc.M710421200
6. Jiang, J.; Tezlaff, C. N.; Takamatsu, S.; Iwatsuki, M.; Komatsu, M.; Ikeda, H.; Cane, D. E. *Biochemistry* **2009,** *48,* 6431–6440. doi:10.1021/bi900766w
7. Dewick, P. M. *Medicinal natural products: a biosynthetic approach;* Wiley: Chichester, 2002.
8. Rohmer, M.; Knani, M.; Simonin, P.; Sutter, B.; Sahm, H. *Biochem. J.* **1993,** *295,* 517–524. doi:10.1042/bj2950517
9. Dickschat, J. S. *Eur. J. Org. Chem.* **2017,** 4872–4882. doi:10.1002/ejoc.201700482
10. Klapschinski, T. A.; Rabe, P.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2016,** *55,* 10141–10144. doi:10.1002/anie.201605425
Angew. Chem. **2016**, *128*, 10296-10299. doi:10.1002/ange.201605425
11. Rabe, P.; Rinkel, J.; Dolja, E.; Schmitz, T.; Nubbemeyer, B.; Luu, T. H.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2017,** *56,* 2776–2779. doi:10.1002/anie.201612439
Angew. Chem. **2017**, *129*, 2820-2823. doi:10.1002/ange.201612439
12. Rinkel, J.; Rabe, P.; Garbeva, P.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2016,** *55,* 13593–13596. doi:10.1002/anie.201608042
Angew. Chem. **2016**, *128*, 13791-13794. doi:10.1002/ange.201608042
13. Korp, J.; Vela Gurovic, M. S.; Nett, M. *Beilstein J. Org. Chem.* **2016,** *12,* 594–607. doi:10.3762/bjoc.12.58
14. Schieferdecker, S.; Domin, N.; Hoffmeier, C.; Bryant, D. A.; Roth, M.; Nett, M. *Eur. J. Org. Chem.* **2015,** 3057–3062. doi:10.1002/ejoc.201500181
15. Braga, D.; Hoffmeister, D.; Nett, M. *Beilstein J. Org. Chem.* **2016,** *12,* 2766–2770. doi:10.3762/bjoc.12.274
16. Kastner, S.; Müller, S.; Natesan, L.; König, G. M.; Guthke, R.; Nett, M. *Arch. Microbiol.* **2012,** *194,* 557–566. doi:10.1007/s00203-012-0789-y
17. Nakano, C.; Oshima, M.; Kurashima, N.; Hoshino, T. *ChemBioChem* **2015,** *16,* 772–781. doi:10.1002/cbic.201402652
18. Takeda, K.; Minato, H.; Ishikawa, M. *Chem. Commun.* **1965,** 79–81. doi:10.1039/c19650000079
19. Takeda, K.; Minato, H.; Ishikawa, M. *Tetrahedron* **1966,** *22* (Suppl. 7), 219–225. doi:10.1016/S0040-4020(01)99108-2
20. Rowe, J. W.; Ronald, R. C.; Nagasampagi, B. A. *Phytochemistry* **1972,** *11,* 365–369. doi:10.1016/S0031-9422(00)90015-3
21. Wratten, S. J.; Faulkner, D. J. *J. Org. Chem.* **1977,** *42,* 3343–3349. doi:10.1021/jo00441a005
22. Piers, E.; Gavai, A. V. *J. Org. Chem.* **1990,** *55,* 2380–2390. doi:10.1021/jo00295a028
23. Bülow, N.; König, W. A. *Phytochemistry* **2000,** *55,* 141–168. doi:10.1016/S0031-9422(00)00266-1

24. Silva, L. F., Jr. *Tetrahedron* **2002,** *58,* 9137–9161. doi:10.1016/S0040-4020(02)00990-0
25. Yamada, Y.; Kuzuyama, T.; Komatsu, M.; Shin-ya, K.; Omura, S.; Cane, D. E.; Ikeda, H. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** *112,* 857–862. doi:10.1073/pnas.1422108112
26. Kashman, Y.; Groweiss, A. *J. Org. Chem.* **1980,** *45,* 3814–3824. doi:10.1021/jo01307a017
27. Schieferdecker, S.; König, S.; Weigel, C.; Dahse, H.-M.; Werz, O.; Nett, M. *Chem. – Eur. J.* **2014,** *20,* 15933–15940. doi:10.1002/chem.201404291

License and Terms

Supporting Information

for

Herpetopanone, a diterpene from *Herpetosiphon aurantiacus* discovered by isotope labeling

Xinli Pan[1,2], Nicole Domin[2], Sebastian Schieferdecker[2], Hirokazu Kage[1], Martin Roth[2] and Markus Nett*[1]

Address: [1]Department of Biochemical and Chemical Engineering, Technical Biology, Technical University Dortmund, Emil-Figge-Strasse 66, 44227 Dortmund, Germany and [2]Leibniz Institute for Natural Product Research and Infection Biology, Hans Knöll Institute, Beutenbergstr. 11a, 07745 Jena, Germany

Email: Markus Nett* - markus.nett@bci.tu-dortmund.de

*Corresponding author

IR and NMR spectra of herpetopanone. Metabolic profile of *H. aurantiacus* 114-95^T

Table of contents

Figure S1: Infrared spectrum of herpetopanone.

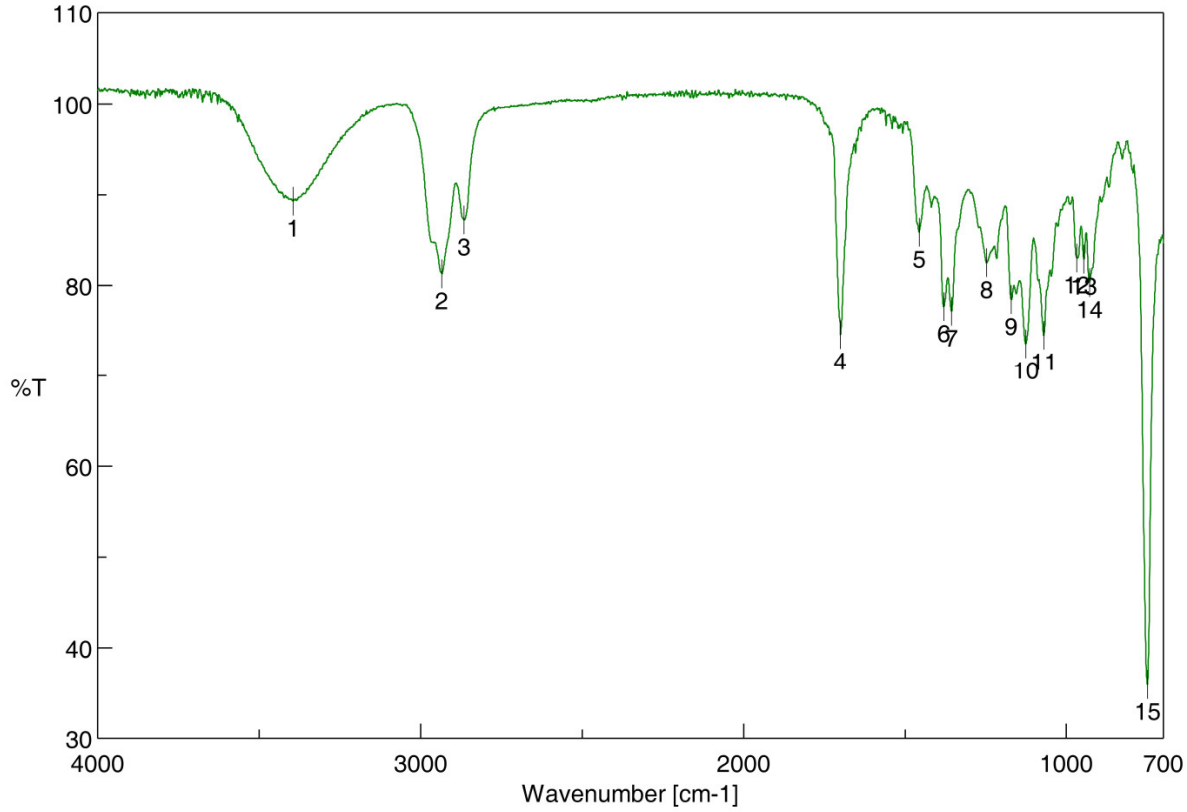


Figure S2: ^{1}H NMR spectrum of herpetopanone in chloroform-d_1.

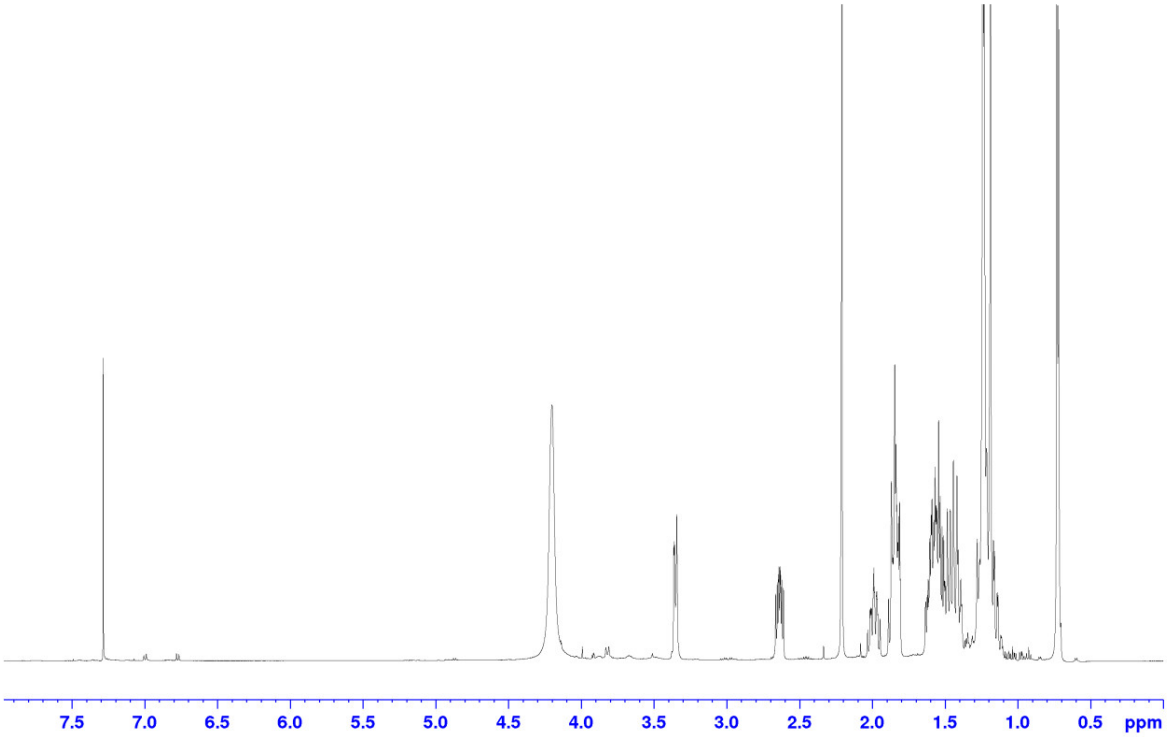


Figure S3: ^{1}H-decoupled ^{13}C NMR spectrum of herpetopanone in chloroform-d_1.

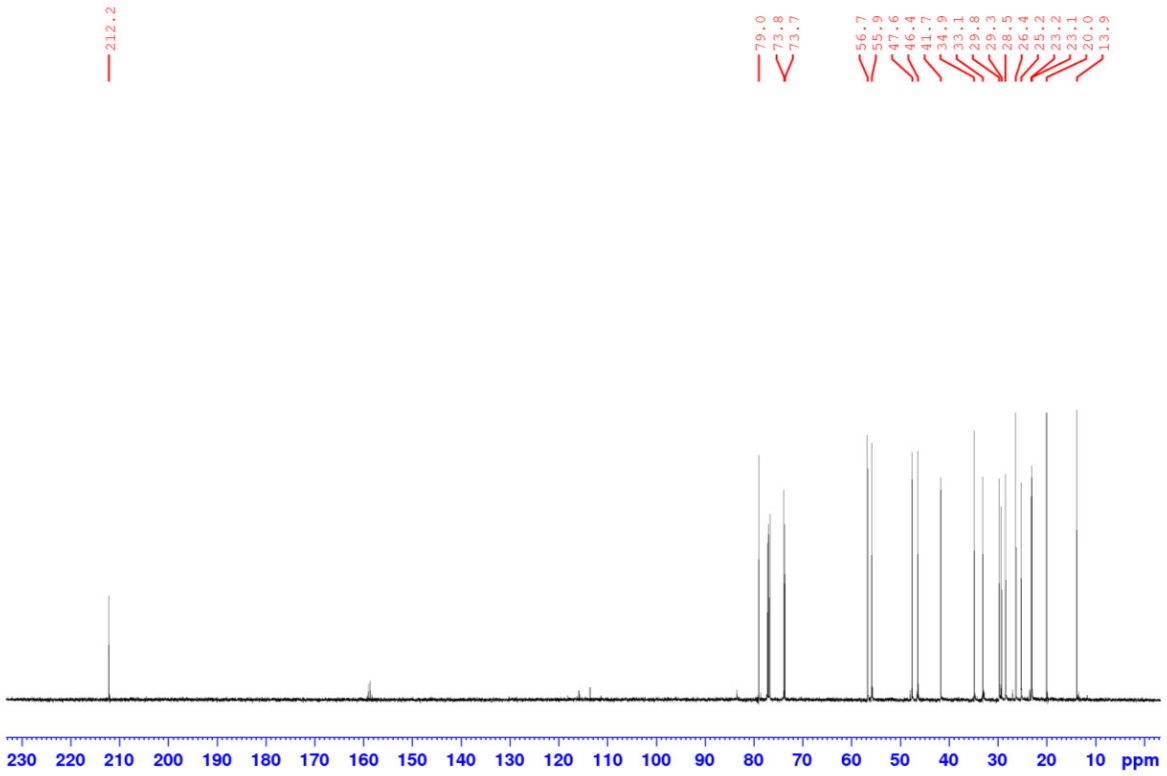


Figure S4: DEPT135 spectrum of herpetopanone in chloroform-d_1.

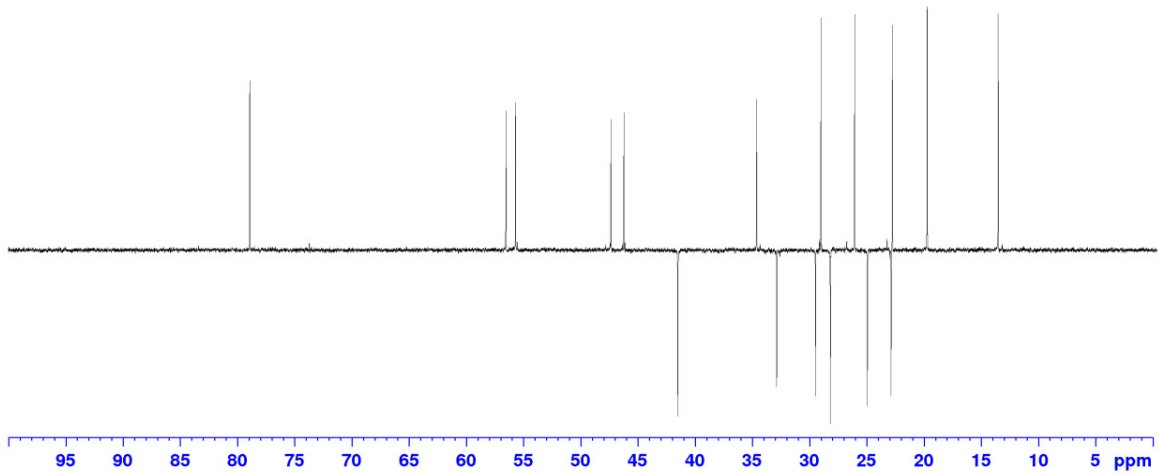


Figure S5: COSY spectrum of herpetopanone in chloroform-d_1.

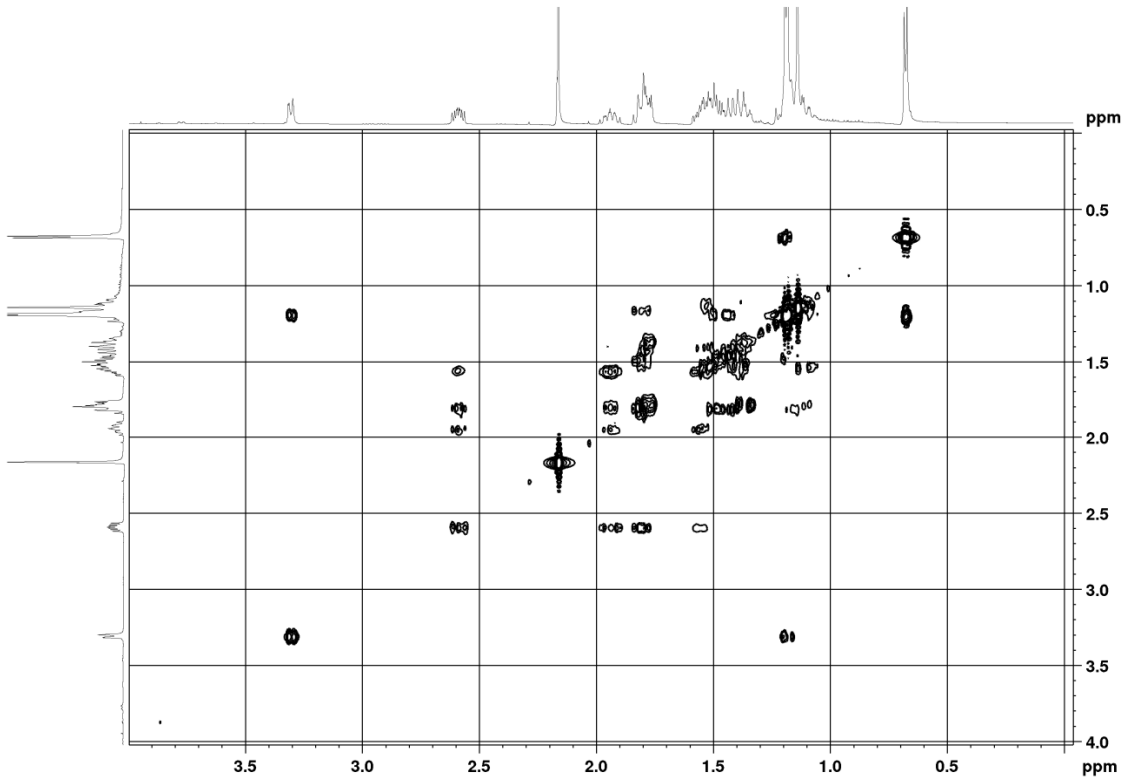


Figure S6: HSQC spectrum of herpetopanone in chloroform-d_1.

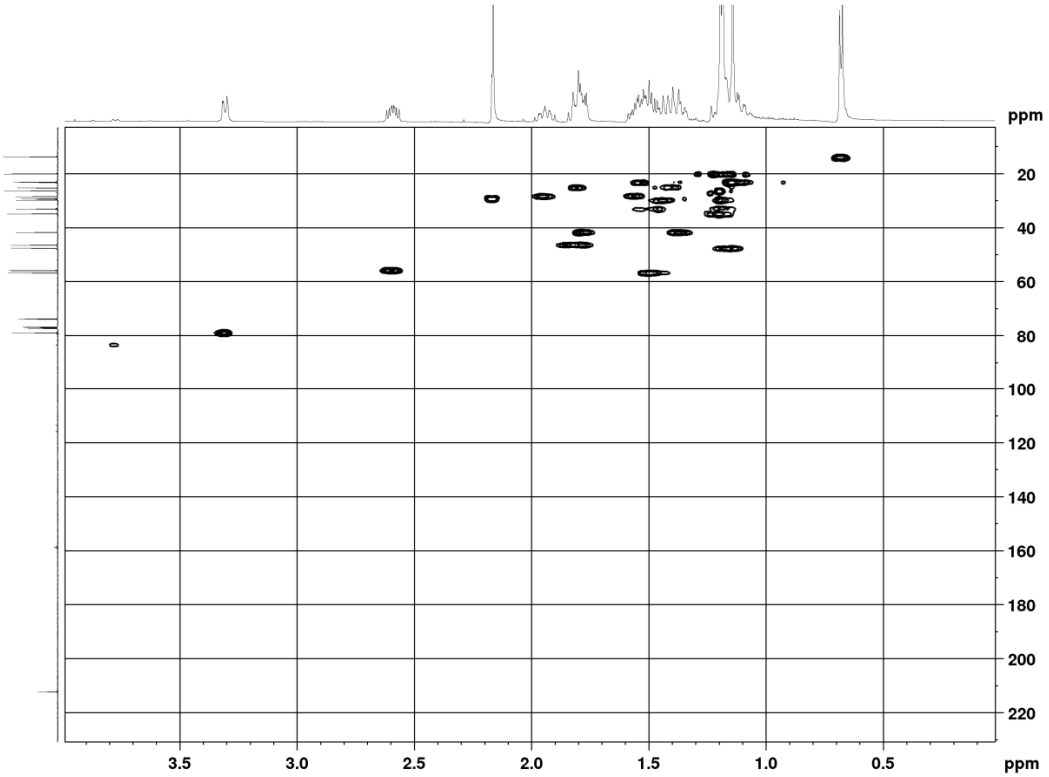


Figure S7: HMBC spectrum of herpetopanone in chloroform-d_1.

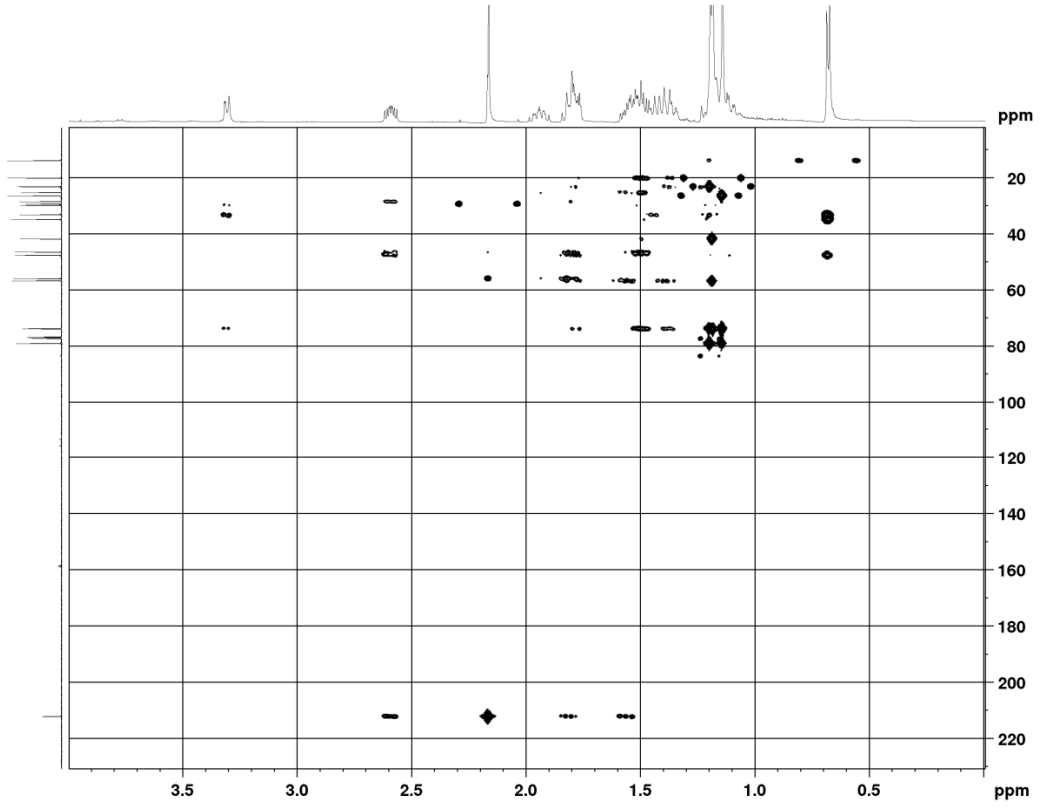


Figure S8: NOESY spectrum of herpetopanone in chloroform-d_1.

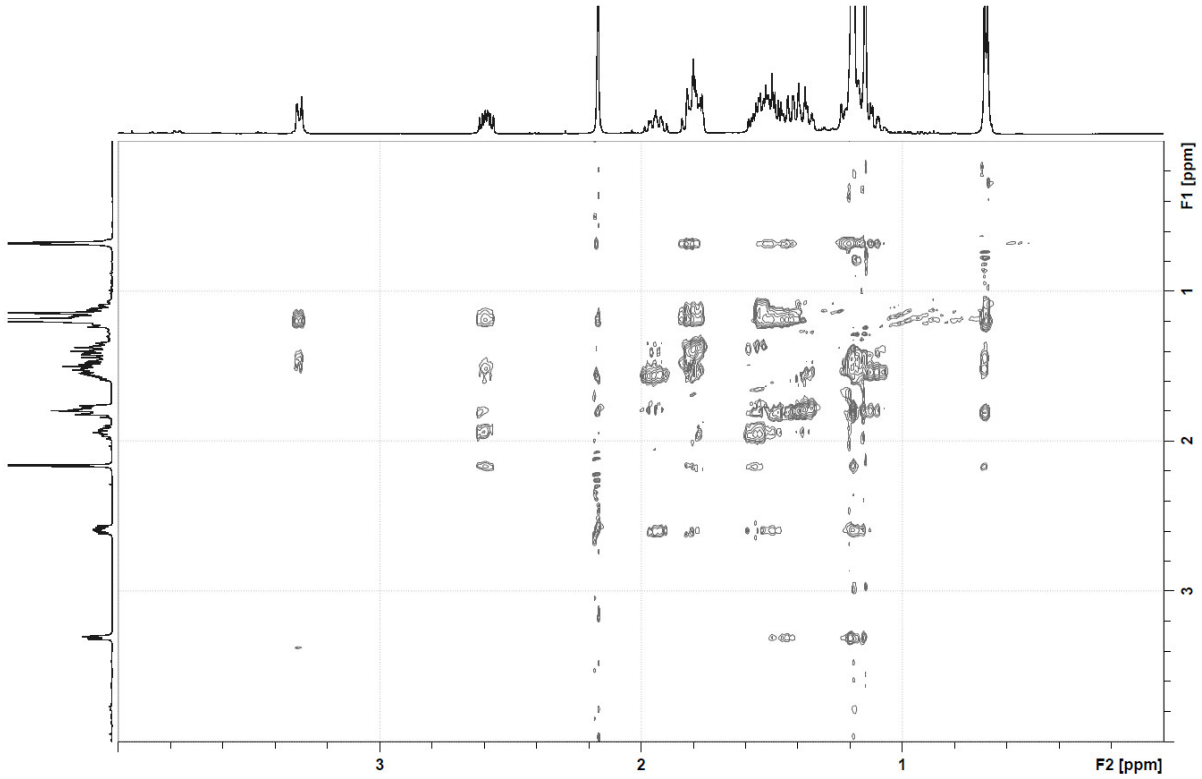
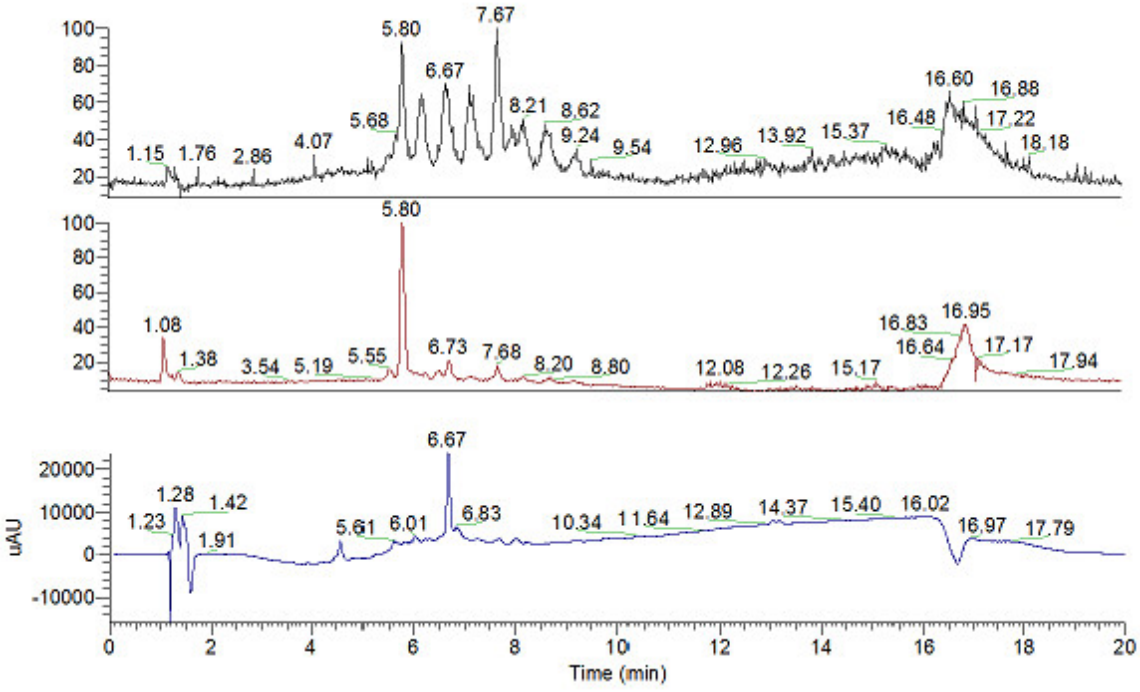


Figure S9: Metabolic profile of *H. aurantiacus* 114-95^T after cultivation in modified VNY medium supplemented with non-labeled D-glucose. Top: Total ion chromatogram (TIC) recorded in positive ion mode. Middle: TIC recorded in negative ion mode. Bottom: UV chromatogram recorded at 254 nm.



3.2 Manuscript B

Xinli Pan, Hirokazu Kage, Karin Martin, and Markus Nett

Herpetosiphon gulosus sp. nov., a filamentous predatory bacterium isolated from sandy soil and *Herpetosiphon giganteus* sp. nov., nom. rev. *Int. J. Syst. Evol. Microbiol.* 2017, 67, 2476-2481

INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

TAXONOMIC DESCRIPTION

Pan *et al.*, *Int J Syst Evol Microbiol* 2017;67:2476–2481

DOI 10.1099/ijsem.0.002141

MICROBIOLOGY SOCIETY

Herpetosiphon gulosus sp. nov., a filamentous predatory bacterium isolated from sandy soil and *Herpetosiphon giganteus* sp. nov., nom. rev.

Xinli Pan,[1,2] Hirokazu Kage,[1] Karin Martin[2,*] and Markus Nett[1,*]

Abstract

Three filamentous gliding bacteria from the German Collection of Microorganisms and Cell Cultures, Hp g11, Hp g471 and Hp g472, were subjected to a phylogenetic analysis. These organisms had previously been classified as members of the genus *Herpetosiphon* based on their growth physiology and morphology. However, a taxonomic assignment at the species level had not been carried out. Analysis of 16S rRNA sequences now confirmed the close relationship of strain Hp g472 to *Herpetosiphon aurantiacus* DSM 785^T (98.6 % nucleotide identity) and *Herpetosiphon geysericola* DSM 7119^T (97.7 %). The results of DNA–DNA hybridization experiments further implied that strain Hp g472 should be classified as a distinct species. The DNA G+C content of strain Hp g472 was 49.9 mol%. The major quinone was MK-10 and the predominant cellular fatty acids were $C_{18:1}$, $C_{16:1}$ and $C_{16:0}$. Based on phenotypic, chemotaxonomic and phylogenetic data it was concluded that strain Hp g472 represents a novel species of the genus *Herpetosiphon*, for which the name *Herpetosiphon gulosus* sp. nov. is proposed. The type strain is Hp g472^T (=DSM 52871^T=NBRC 112829^T). In contrast to Hp g472^T, the strains Hp g11 and Hp g471 exhibited closest 16S rRNA gene sequence similarity (>99 %) with '*Herpetosiphon giganteus*' Hp a2. The distinctive genotypic and phenotypic properties of the latter supported the revival of the name as *Herpetosiphon giganteus* (*ex* Reichenbach & Golecki, 1975) sp. nov., nom. rev. We propose the previously deposited reference strain DSM 589^T=NBRC 112828^T as the type strain.

Herpetosiphon is the type genus of the family *Herpetosiphonaceae* and was first described in 1968 [1]. Phylogenetic studies established its taxonomic position in the green non-sulfur bacteria [2], which later became the phylum *Chloroflexi* [3]. At the time of writing, there are only two species in the genus *Herpetosiphon* with validly published names: *Herpetosiphon aurantiacus* and *Herpetosiphon geysericola* isolated from soil, freshwater, and decaying organic matter [4–6]. They are aerobic chemoheterotrophs, and glide over solid surfaces [3]. A consistent feature not only of *Herpetosiphon* spp., but also of other members of the *Chloroflexi* is their unusual cell wall composition. These bacteria stain Gram-negative, but lack a lipopolysaccharide-containing outer membrane. Furthermore, their peptidoglycan complex contains L-ornithine as the diamino acid instead of the more common *meso*-diaminopimelic acid [7].

Several strains of species of the genus *Herpetosiphon* have been reported to feed on other bacteria [8, 9], but it is unclear whether predatory behavior is a common trait in this genus. For the killing of other microbes, *Herpetosiphon* spp. are assumed to resort to a wolf pack strategy, in which a large number of predatory cells congregate in order to lyse the prey by the combined secretion of hydrolytic enzymes [10]. Recent studies suggest further that strains of species of the genus *Herpetosiphon* are competent producers of bioactive secondary metabolites [11–15]. The biosynthetic potential of the genus is supported by genomic analyses of *H. aurantiacus* and *H. geysericola* [16, 17]. It has even been speculated that a correlation between predation and antibiotic biosynthesis might exist, as previously demonstrated for myxobacteria [18–20]. This would make the genus *Herpetosiphon* an interesting resource for drug discovery.

A large collection of about 460 *Herpetosiphon* and *Herpetosiphon*-like strains is currently available at the German Collection of Microorganisms and Cell Cultures (DSMZ), but only few representatives of this collection have been assigned down to the species level [4]. Considering the lack of studies in the field, as well as the possible

Author affiliations: [1]Department of Biochemical and Chemical Engineering, Technical University Dortmund, Emil-Figge-Str. 66, 44227 Dortmund, Germany; [2]Leibniz Institute for Natural Product Research and Infection Biology e. V., Hans-Knöll- Institute, Beutenbergstr. 11a, 07745 Jena, Germany.
***Correspondence:** Karin Martin, karin.martin@leibniz-hki.de; Markus Nett, markus.nett@bci.tu-dortmund.de
Keywords: *Herpetosiphon*; *Herpetosiphon gulosus*; *Herpetosiphon giganteus*; predatory bacteria.
Abbreviation: DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen (German Collection of Microorganisms and Cell Cultures).
The GenBank/EMBL accession numbers for the partial 16S rRNA sequences of strains Hp a2, Hp g11, Hp g471 and Hp g472 are KY689830, KY689831, KY689832, and KY689833, respectively.

biotechnological potential of these bacteria, we decided to investigate the taxonomic standing of three, as-yet, unassigned strains of species of the genus *Herpetosiphon* from the DSMZ collection. Strain Hp g11 (DSM 52868) had been isolated from a forest soil collected near Wolfenbuettel, Germany. Strain Hp g471 (DSM 52870) originated from a soil sample taken in San Diego, USA, whereas strain Hp g472 (DSM 52871) had been isolated from a sandy soil sample collected from the beach of Poel island, Germany. Furthermore, we included strain Hp a2 (DSM 589) in our study. This bacterium, which had been retrieved from a Nepalese soil sample, had originally been classified as '*Herpetosiphon giganteus*'. The assignment was made together with the proposal to abandon the species epithet *aurantiacus* in favor of *giganteus* [21]. Later, the strain was designated as *Herpetosiphon aurantiacus* Hp a2 [22]. To our knowledge, however, strain Hp a2 has never been subjected to a phylogenetic or chemotaxonomic analysis. It was hence unclear whether '*H. giganteus*' and *H. aurantiacus* actually represented the same species.

For experimental analyses, the four strains were grown at 30 °C on VY/2 agar [0.5 % (w/v) baker's yeast (fresh weight), 0.1 % (w/v) $CaCl_2.2H_2O$, 1.5 % (w/v) agar (pH 7.2); supplemented with 0.5 mg l^{-1} vitamin B12], on CY agar [0.3 % (w/v) casitone, 0.3 % (w/v) yeast extract, 0.1 % (w/v) $CaCl_2.2H_2O$, 1.5 % (w/v) agar (pH 7.2)], or in modified Pol medium [0.3 % (w/v) starch, 0.3 % (w/v) casitone, 0.05 % (w/v) $CaCl_2.2H_2O$, 0.2 % (w/v) $MgSO_4.7H_2O$, 1.19 % (w/v) HEPES (pH 7.2); supplemented with 1 ml l^{-1} trace elements solution (0.01 % (w/v) $MnCl_2.4H_2O$, 0.002 % (w/v) $CoCl_2$, 0.001 % (w/v) CuSO4, 0.001 % (w/v) $Na_2MoO_4.2H_2O$, 0.002 % (w/v) $ZnCl_2$, 0.0005 % (w/v) LiCl, 0.0005 % (w/v) $SnCl_2.2H_2O$, 0.001 % (w/v) H_3BO_3, 0.002 % (w/v) KBr, 0.002 % (w/v) KI, 0.0008 % (w/v) EDTA]. Cell morphology was examined after growth on CY agar using a Zeiss Axio-Imager microscope. For scanning electron microscopy, cells from a 7-day-old submers culture were fixed with 0.5 % (w/v) glutaraldehyde, washed and dehydrated using increasing concentrations of ethanol. After sputter coating with gold-palladium, the cells were examined with a Zeiss model 962 scanning electron microscope. The composition of cellular fatty acids was determined according to Sasser [23] and Cole *et al.* [24]. Menaquinones were extracted as described by Collins *et al.* [25] and analysed by HPLC. Enzyme activities were determined using conventional test panels, including API 50 CH, API ZYM and API 20 NE. For this, the strains of species of the genus *Herpetosiphon* were grown in modified Pol medium at 30 °C for 7 days. Effects of pH, temperature and NaCl concentration were assessed using VY/2 agar cultures. Each test was run in triplicate and the outcome was evaluated after incubation for 7 days. To determine the pH range for growth, the medium had been adjusted to pH 4.0–9.0 and incubation was conducted at 30 °C. Temperature tolerance was tested at 4, 20, 26, 30, 37 and 45 °C (pH 7.2). Halotolerance was evaluated with varying concentrations of NaCl, [1, 2, 3, 4 or 5 % (w/v)]. Susceptibility to antibiotics was tested by placing antibiotic discs (BBL Sensi-Disc, Becton Dickinson) on VY/2 agar plates seeded with suspensions of the tested strains, which were grown for 3 to 4 days at 28 °C. Predatory activity was tested in a lawn predation assay. Prey bacteria (*Escherichia coli* DSM 18039, *Chromobacterium pseudoviolaceum* DSM 23279^T, *Bacillus subtilis* 168 and *Rhodococcus rhodochrous*

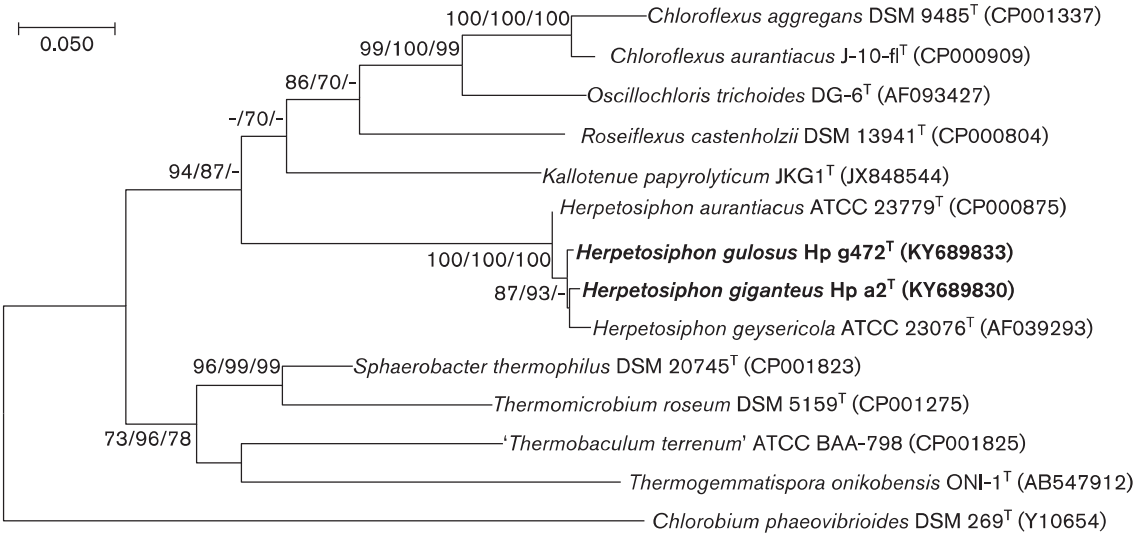


Fig. 1. Phylogenetic tree inferred from 16S rRNA gene sequences, showing the phylogenetic positions of strains Hp g472 and Hp a2 relative to the type strains of species of the genus *Herpetosiphon* and further representatives of the phylum *Chloroflexi*. Support values from maximum-likelihood/neighbour-joining/maximum-parsimony bootstrapping are shown above the branches if equal to or larger than 70 %. Bar, 0.05 substitutions per site.

DSM 43334) were spotted as suspensions (500 µl, 10^8 cells ml^{-1}) on TPM agar plates (10 mM Tris-HCl, 1 mM KH_2PO_4, 8 mM $MgSO_4$, 1.5 % (w/v) agar, pH 7.6). Subsequently, the prey-covered plates were individually inoculated with the different strains of species of the genus *Herpetosiphon* following the procedure described by Seccareccia *et al.* [26]. A prey-free agar plate served as a control. Every experiment was repeated three times. The diameter of the swarm was measured on days 1 and 7.

For nucleic acid studies, genomic DNA was isolated following cell lysis using lysozyme [27]. The DNA was purified from the resulting crude lysate via column chromatography on hydroxyapatite [28]. The DNA G+C content was determined by HPLC after digestion with nuclease P_1 [29]. DNA–DNA hybridization experiments were performed in duplicate using the methods of De Ley *et al.* [30] and Huss *et al.* [31]. The nearly full-length 16S rRNA gene sequences of the four strains were obtained after PCR amplification using primers 27 f and 1492 r [32]. All PCR products were purified with a QIAquick PCR purification kit (Qiagen) prior to sequencing. For phylogenetic analysis, the 16S rRNA gene sequences of the type strains *H. aurantiacus* DSM 785 and *H. geysericola* DSM 7119 were retrieved from GenBank. Furthermore, representatives of the closely related genera *Chloroflexus*, *Oscillochloris*, *Roseiflexus*, *Kallotenue*, '*Thermobaculum*', *Thermogemmatispora*, *Thermomicrobium* and *Sphaerobacter* were included in the analysis. The 16S rRNA sequence of *Chlorobium phaeovibrioides* DSM 269^T was used as the outgroup. Sequences were aligned by using CLUSTAL X [33]. The 16S rRNA gene-based tree (Fig. 1) was reconstructed with the MEGA7 program package [34]. To estimate the confidence of the tree topologies, bootstrap-resampling analysis with 1000 replicates was performed for neighbour-joining [35], maximum-likelihood [36] and maximum-parsimony methods [37]. Pairwise sequence similarities were determined using the Calculate Matrix function of ARB software with Jukes and Cantor correction [38, 39].

Comparison of 16S rRNA gene sequences confirmed that strains Hp g11, Hp g471, Hp g472, and Hp a2 are closely related to the two type strains of species of the genus *Herpetosiphon*. Strain Hp a2 possessed a 16S rRNA sequence similarity of 98.3 % to *H. aurantiacus* DSM 785^T and of 97.6 % to *H. geysericola* DSM 7119^T. In the case of strain Hp g472, the respective values were 98.6 and 97.7 %. The sequence similarity between strains Hp a2 and Hp g472 was 98.6 %. Strains Hp g11 and Hp g471 shared almost identical 16S rRNA gene sequences with Hp a2. Due to their high homology, only Hp a2 and Hp g472 were included in the following characterization. Strain Hp g472 exhibited low DNA–DNA reassociation values with *H. aurantiacus* DSM 785^T (37.9 %), *H. geysericola* DSM 7119^T (29 %) and strain Hp a2 (18.7 %). The DNA–DNA relatedness values of strain Hp a2 to *H. aurantiacus* DSM 785^T and *H. geysericola* DSM 7119^T were 36.2 and 68.9 %, respectively. These data, hence, indicate that strains Hp a2 and Hp g472 represent distinct species.

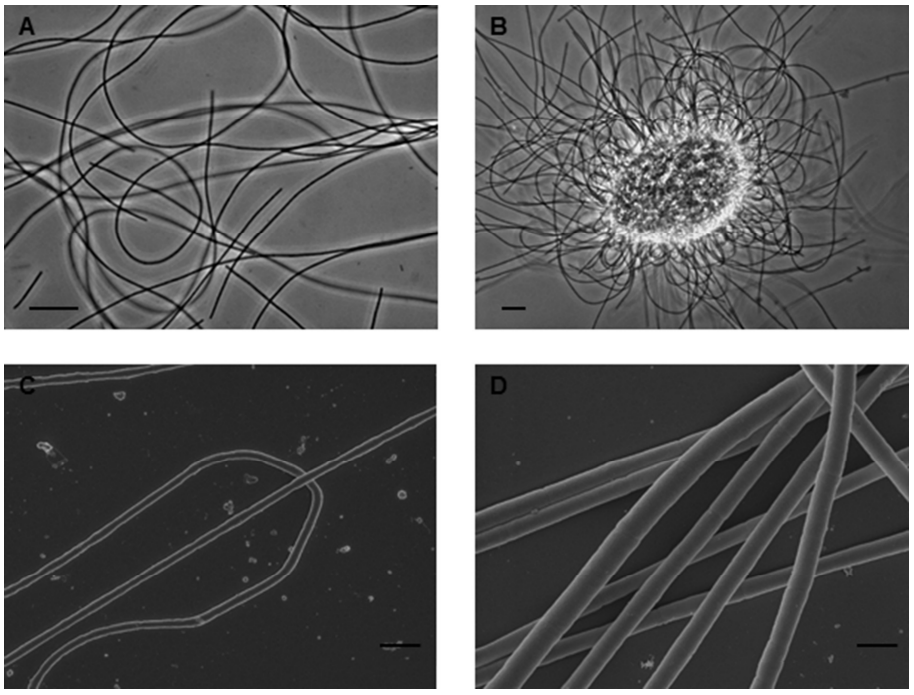


Fig. 2. Morphology of species of the genus *Herpetosiphon*. Filaments of strains Hp g472^T (A) and Hp a2^T (B) under a phase contrast microscope (bars=20 µm). Scanning electron micrographs of filaments from strains Hp g472^T (C, bar=2 µm) and Hp a2^T (D, bar=1 µm).

Table 1. Phenotypic characteristics of strains Hp g472^T and Hp a2^T in comparison to the type strains of closely related species of the genus *Herpetosiphon*

Strains: 1, *Herpetosiphon gulosus* sp. nov. DSM 52871^T (Hp g472^T); 2, *H. giganteus* sp. nov., nom. rev. DSM 589^T (Hp a2^T); 3, *H. aurantiacus* DSM 785^T; 4, *H. geysericola* DSM 7119^T. +, Positive; ±, weakly positive; −, negative.

Characteristic	1	2	3	4
pH for growth	6–9	6–9	6–9	6–9
Temperature for growth (°C)	20–37	20–37	20–37	20–37
NaCl tolerance (%)	2	3	1	0
Production of acid from:				
N-Acetylglucosamine	−	−	+	−
Aesculin	+	+	−	+
Glycogen	±	−	−	−
Hydrolysis of:				
Arginine	+	+	+	−
Gelatin	−	+	−	−
Glucose fermentation	+	±	±	−
Enzyme activities:				
Acid phosphatase	+	−	+	+
Alkaline phosphatase	+	+	+	−
Esterase (C4)	−	±	±	±
Esterase lipase (C8)	−	±	±	+
α-Galactosidase	±	−	+	+
Cysteine arylamidase	−	−	±	−
Leucine arylamidase	−	−	+	−
Valine arylamidase	−	−	±	−
Trypsin	−	−	+	+
Resistance to:				
Ampicillin + sulbactam (10+10 µg)	−	−	+	+
Tetracyclin (30 µg)	+	−	−	−
Nalidixic acid (30 µg)	+	−	−	+
Azlocillin (75 µg)	−	−	−	+
Amoxicillin + clavulanic acid (20+10 µg)	−	−	−	+
Lincomycin (2 µg)	+	+	−	+
Sulfonamide (200 µg)	+	−	−	−
Cefoxitin (30 µg)	+	−	+	+
Gentamicin (10 µg)	−	−	+	+

The DNA G+C content of strains Hp g472 and Hp a2 was 49.9 mol% and 52.6 mol%, respectively. Both values are consistent with those observed for species of the genus *Herpetosiphon* [4]. The morphological features of Hp a2 and Hp g472 were also in accordance with their classification as members of the genus *Herpetosiphon*. Both bacteria formed unbranched filaments with a diameter ranging from 0.7 to 1.2 µm (Fig. 2). Filaments were flexible and often more than 100 µm in length. Characteristic translucent segments were frequently observed at the ends of the filaments. Temperature and pH requirements for growth were consistent among all strains of species of the genus *Herpetosiphon*. In contrast, the extent of halotolerance varied. While *H. geysericola* did not even grow at the lowest NaCl concentration tested (1 %), the other strains tolerated up to 3 % (w/v) NaCl. The carbohydrate fermentation patterns of the strains were largely identical, except for the utilization of N-acetylglucosamine, which was a distinctive feature of *H. aurantiacus*. On the other hand, Hp g472 and Hp a2 could be easily distinguished from each other and the two type strains by their unique profiles of enzymic activities, which are summarized in Table 1.

The major fatty acids of strain Hp g472 were $C_{18:1}$ (49.8 %), $C_{16:1}$ (20.5 %) and $C_{16:0}$ (16.7 %). For *H. giganteus* Hp a2, fatty acid analysis revealed the same principal components, albeit with slight differences in the respective content: $C_{18:1}$ (34.8 %), $C_{16:1}$ (13.4 %), and $C_{16:0}$ (10.9 %). The predominance of straight-chain fatty acids in *Herpetosiphon* spp. has previously been reported [4]. The predominant menaquinone of both strains was MK-10. Furthermore, strain Hp a2 exhibited minor amounts of MK-9 (MK-10: MK-9=81 : 11), while strain Hp g472 contained small amounts of MK-9, as well as MK-6 (MK-10: MK-9: MK-6=81 : 13 : 4).

The swarming of the different strains of species of the genus *Herpetosiphon* on lawns of living bacteria also involved the lysis of the latter. From this obervation, it was concluded that all strains of species of the genus *Herpetosiphon* tested were capable of predation. Interestingly, the swarming rates differed considerably among the four strains (Fig. 3). When averaged across all prey, strain Hp g472 showed the strongest predatory swarming, followed by *H. geysericola* DSM 7119^T, strain Hp a2 and *H. aurantiacus* DSM 785^T. Similar to reports from other predatory bacteria [26, 40], the strains of species of the genus *Herpetosiphon* also displayed unique prey preferences. Thus, strain Hp g472 swarmed faster on *E. coli* than on the other prey bacteria. In contrast, strain Hp a2 showed its largest swarm expansion on *B. subtilis*. Contrary to the report of Quinn and Skerman [9], we also observed feeding of members of the genus *Herpetosiphon* on *E. coli*. A possible explanation for this apparent inconsistency is the absence of organic nutrients other than living prey bacteria in the TPM agar used in this study, while the assay medium used by Quinn and Skerman contained low concentrations of peptone and yeast extract [9].

In summary, the 16S rRNA-derived phylogenetic grouping confirmed that strains Hp g 472 and Hp a2 are members of the genus *Herpetosiphon*. Due to the high 16S rRNA sequence homology with closely related strains, however, DNA–DNA hybridization experiments were necessary to resolve the species status of the two strains. These analyses indicate that both strains should be assigned to distinct species. Several phenotypic characteristics, including the predatory performance, support a distinctiveness from the two species of the genus *Herpetosiphon*. Consolidating the genotypic and phenotypic data, we propose that strain Hp g472 represents a novel species of the genus of *Herpetosiphon*, for which the name *Herpetosiphon gulosus* sp. nov. is proposed. Based on a review of the literature and new data obtained in this study, the name *Herpetosiphon giganteus* (sp. nov., nom. rev.) is revived for strain Hp a2.

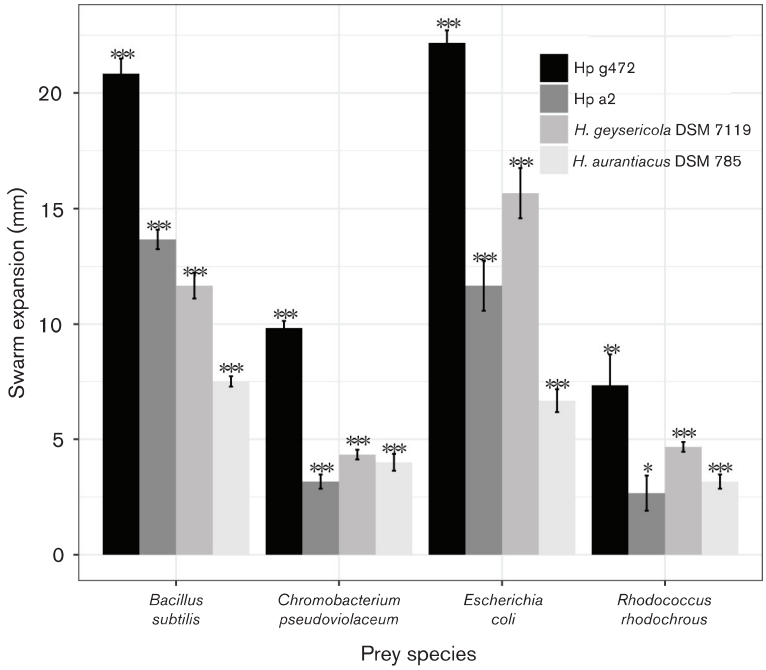


Fig. 3. Effect of strains of species of the genus *Herpetosiphon* on different prey bacteria, as determined by the lawn predation assay. The mean swarm expansion is shown (±95 % confidence interval [n=3]). Paired *t* test: *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$ between day 1 and day 7.

DESCRIPTION OF *HERPETOSIPHON GIGANTEUS* (*EX* REICHENBACH AND GOLECKI, 1975) SP. NOV., NOM. REV

Herpetosiphon giganteus (gi.gan′ te.us. L. masc. adj. *giganteus* giant, gigantic).

Characteristics are as described by Reichenbach and Golecki [21] with the following emendation. Growth occurs under aerobic conditions at 20–37 °C and pH 6–9. Tolerates up to 3 % (w/v) NaCl. Cells can utilize L-arabinose, starch, sucrose, galactose, L-rhamnose, glycerol, raffinose, glucose, maltose, D-xylose, fructose, trehalose, lactose and mannitol as the sole carbon sources for energy. Acid is produced from aesculin. Cells exhibit predatory activity in the absence of alternative nutrient sources. Cells are resistant to penicillin G, kanamycin, polymyxin B, methicillin, nystatin, carbenicillin and lincomycin. Cellular fatty acids are mainly composed of $C_{18:1}$, $C_{16:0}$ and $C_{16:1}$. The major quinone is MK-10. The DNA G+C content is 52.6 mol%.

The type strain is Hp a2^T (=DSM 589^T=NBRC 112828^T), isolated from a Nepalese soil sample.

DESCRIPTION OF *HERPETOSIPHON GULOSUS* SP. NOV.

Herpetosiphon gulosus (gu.lo′ sus. L. masc. adj. *gulosus* gluttonous, pertaining to the strong predatory activity of the type strain).

Gram-stain-negative bacterium forms unbranched, multicellular filaments of varying length with a diameter of 0.7–1.2 µm. Growth occurs under aerobic conditions at 20–37 °C, pH 6–9 and a maximum NaCl concentration of 2 % (w/v). Utilizes L-arabinose, starch, sucrose, galactose, L-rhamnose, glycerol, raffinose, glucose, maltose, D-xylose, fructose, trehalose, lactose and mannitol as the sole carbon sources for energy. Acid is produced from glycogen and aesculin. Cells exhibit strong predatory activity in the absence of alternative nutrient sources. Cells are resistant to penicillin G, kanamycin, polymyxin B, methicillin, nystatin, carbenicillin, tetracyclin, nalidixic acid, lincomycin, sulfonamide and cefoxitin. The major cellular fatty acids are $C_{18:1}$, $C_{16:0}$ and $C_{16:1}$. The predominant menaquinone is MK-10. The DNA G+C content is 49.9 %.

The type strain is Hp g472^T (=DSM 52871^T=NBRC 112829^T), isolated from sand and plant residues collected from the beach of Poel island, Germany.

Funding information
This work received no specific grant from any funding agency.

Acknowledgements
X. Pan thanks the Chinese Scholarship Council (CSC) for a doctoral stipend. The authors are indebted to Susanne Verbarg (Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures) for providing the strains and expert advice. The authors are also grateful to Karin Burmeister and Bettina Bardl (Leibniz Institute for Natural Product Research and Infection Biology) for excellent technical assistance.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

1. **Holt JG, Lewin RA.** *Herpetosiphon aurantiacus* gen. et sp. n., a new filamentous gliding organism. *J Bacteriol* 1968;95:2407–2408.
2. **Oyaizu H, Debrunner-Vossbrinck B, Mandelco L, Studier JA, Woese CR.** The green non-sulfur bacteria: a deep branching in the eubacterial line of descent. *Syst Appl Microbiol* 1987;9:47–53.
3. **Garrity GM, Holt JG, Phylum BVI.** *Chloroflexi* phy. nov. In: Boone DR, Castenholz RW and Garrity GM (editors). *Bergey's Manual of Systematic Bacteriology*, 2nd ed, vol. 1, The Archaea and the Deeply Branching and Phototrophic Bacteria. New York: Springer; 2001. pp. 427–446.
4. **Lee N, Reichenbach H.** The genus *Herpetosiphon*. In: Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E *et al.* (editors). *The Prokaryotes,*. New York: Springer; 2006. pp. 854–877.
5. **Euzéby JP.** List of bacterial names with standing in nomenclature: a folder available on the internet. *Int J Syst Bacteriol* 1997;47:590–592.
6. **Sly LI, Taghavi M, Fegan M.** Phylogenetic heterogeneity within the genus *Herpetosiphon*: transfer of the marine species *Herpetosiphon cohaerens*, *Herpetosiphon nigricans* and *Herpetosiphon* persicus to the genus *Lewinella* gen. nov. in the *Flexibacter-Bacteroides-Cytophaga* phylum. *Int J Syst Bacteriol* 1998;48 Pt 3:731–737.
7. **Jürgens UJ, Meissner J, Reichenbach H, Weckesser J.** L-ornithine containing peptidoglycan-polysaccharide complex from the cell wall of the gliding bacterium *Herpetosiphon aurantiacus*. *FEMS Microbiol Lett* 1989;60:247–250.
8. **Lewin RA.** New *Herpetosiphon* species (*Flexibacterales*). *Can J Microbiol* 1970;16:517–520.
9. **Quinn GR, Skerman VBD.** *Herpetosiphon*—Nature's scavenger? *Curr Microbiol* 1980;4:57–62.
10. **Jurkevitch E.** Predatory behaviors in bacteria — diversity and transitions. *Microbe Magazine* 2007;2:67–73.
11. **Kastner S, Müller S, Natesan L, König GM, Guthke R *et al.*** 4-Hydroxyphenylglycine biosynthesis in *Herpetosiphon aurantiacus*: a case of gene duplication and catalytic divergence. *Arch Microbiol* 2012;194:557–566.
12. **Mir Mohseni M, Höver T, Barra L, Kaiser M, Dorrestein PC *et al.*** Discovery of a mosaic-like biosynthetic assembly line with a decarboxylative off-loading mechanism through a combination of genome mining and imaging. *Angew Chem Int Ed Engl* 2016;55:13611–13614.
13. **Nakano C, Oshima M, Kurashima N, Hoshino T.** Identification of a new diterpene biosynthetic gene cluster that produces O-methylkolavelool in *Herpetosiphon aurantiacus*. *Chembiochem* 2015;16:772–781.
14. **Nett M, Erol O, Kehraus S, Köck M, Krick A *et al.*** Siphonazole, an unusual metabolite from *Herpetosiphon* sp. *Angew Chem Int Ed Engl* 2006;45:3863–3867.
15. **Schieferdecker S, Domin N, Hoffmeier C, Bryant DA, Roth M *et al.*** Structure and absolute configuration of auriculamide, a natural product from the predatory bacterium *Herpetosiphon aurantiacus*. *Eur J Org Chem* 2015;14:3057–3062.
16. **Kiss H, Nett M, Domin N, Martin K, Maresca JA *et al.*** Complete genome sequence of the filamentous gliding predatory bacterium *Herpetosiphon aurantiacus* type strain (114-95T). *Stand Genomic Sci* 2011;5:356–370.
17. **Ward LM, Hemp J, Pace LA, Fischer WW.** Draft genome sequence of *Herpetosiphon geysericola* GC-42, a nonphototrophic member of the *Chloroflexi* class *Chloroflexia*. *Genome Announc* 2015;3:e01352-15–15.
18. **Korp J, Vela Gurovic MS, Nett M.** Antibiotics from predatory bacteria. *Beilstein J Org Chem* 2016;12:594–607.
19. **Schieferdecker S, König S, Weigel C, Dahse HM, Werz O *et al.*** Structure and biosynthetic assembly of gulmirecins, macrolide antibiotics from the predatory bacterium *Pyxidicoccus fallax*. *Chemistry* 2014;20:15933–15940.
20. **Xiao Y, Wei X, Ebright R, Wall D.** Antibiotic production by myxobacteria plays a role in predation. *J Bacteriol* 2011;193:4626–4633.
21. **Reichenbach H, Golecki JR.** The fine structure of *Herpetosiphon*, and a note on the taxonomy of the genus. *Arch Microbiol* 1975;102:281–291.
22. **Harwardt R, Maier E, Reichenbach H, Weckesser J, Benz R.** Channel-forming (Porin) activity in *Herpetosiphon aurantiacus* Hp a2. *J Bacteriol* 2004;186:6667–6670.
23. **Sasser M.** Bacterial identification by gas chromatographic analysis of fatty acid methyl esters GC-FAME. *MIDI Technical Note* 2006.
24. **Cole JK, Gieler BA, Heisler DL, Palisoc MM, Williams AJ *et al.*** *Kallotenue papyrolyticum* gen. nov., sp. nov., a cellulolytic and filamentous thermophile that represents a novel lineage (*Kallotenuales* ord. nov., *Kallotenuaceae* fam. nov.) within the class *Chloroflexia*. *Int J Syst Evol Microbiol* 2013;63:4675–4682.
25. **Collins MD, Pirouz T, Goodfellow M, Minnikin DE.** Distribution of menaquinones in actinomycetes and corynebacteria. *J Gen Microbiol* 1977;100:221–230.
26. **Seccareccia I, Kost C, Nett M.** Quantitative analysis of *Lysobacter* predation. *Appl Environ Microbiol* 2015;81:7098–7105.
27. **Marmur J.** A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *J Mol Biol* 1961;3:208–IN1.
28. **Cashion P, Holder-Franklin MA, Mccully J, Franklin M.** A rapid method for the base ratio determination of bacterial DNA. *Anal Biochem* 1977;81:461–466.
29. **Mesbah M, Premachandran U, Whitman WB.** Precise measurement of the G+C content of deoxyribonucleic acid by high-performance liquid chromatography. *Int J Syst Bacteriol* 1989;39:159–167.
30. **de Ley J, Cattoir H, Reynaerts A.** The quantitative measurement of DNA hybridization from renaturation rates. *Eur J Biochem* 1970;12:133–142.
31. **Huss VA, Festl H, Schleifer KH.** Studies on the spectrophotometric determination of DNA hybridization from renaturation rates. *Syst Appl Microbiol* 1983;4:184–192.
32. **Lane DJ.** Nucleic acid techniques in bacterial systematics. In: Stackebrandt E and Goodfellow M (editors). *Nucleic Acid Techniques in Bacterial Systematics*. Chichester: Wiley; 1991. pp. 115–174.
33. **Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA *et al.*** CLUSTAL W and CLUSTAL X version 2.0. *Bioinformatics* 2007;23:2947–2948.
34. **Kumar S, Stecher G, Tamura K.** MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016;33:1870–1874.
35. **Saitou N, Nei M.** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987;4:406–425.
36. **Guindon S, Delsuc F, Dufayard JF, Gascuel O.** Estimating maximum likelihood phylogenies with PhyML. *Methods Mol Biol* 2009;537:113–137.
37. **Takahashi K, Nei M.** Efficiencies of fast algorithms of phylogenetic inference under the criteria of maximum parsimony, minimum evolution, and maximum likelihood when a large number of sequences are used. *Mol Biol Evol* 2000;17:1251–1258.
38. **Jukes TH, Cantor CR.** Evolution of protein molecules. In: Munro HN (editor). *Mammalian Protein Metabolism*. New York: Academic Press; 1969. pp. 21–132.
39. **Ludwig W, Strunk O, Westram R, Richter L, Meier H *et al.*** ARB: a software environment for sequence data. *Nucleic Acids Res* 2004;32:1363–1371.
40. **Morgan AD, Maclean RC, Hillesland KL, Velicer GJ.** Comparative analysis of *Myxococcus* predation on soil bacteria. *Appl Environ Microbiol* 2010;76:6920–6927.

3.3 Manuscript C

Xinli Pan and Markus Nett

Natural product biosynthesis in the phylum *Chloroflexi*: a genomics perspective.

(Manuscript in preparation)

Natural product biosynthesis in the phylum *Chloroflexi*: a genomics perspective

Xinli Pan[1,2] and Markus Nett*[1]

Address: [1]Department of Biochemical and Chemical Engineering, Technical Biology, Technical University Dortmund, Emil-Figge-Strasse 66, 44227 Dortmund, Germany and [2]Leibniz Institute for Natural Product Research and Infection Biology – Hans-Knöll-Institute, Beutenbergstr. 11, 07745 Jena, Germany

Email: Markus Nett* - markus.nett@bci.tu-dortmund.de

* Corresponding author

Keywords: genome mining, *Chloroflexi*, *Herpetosiphon*, *Ktedonobacter*, *Thermogemmatispora,* biosynthesis, gene clusters.

Abstract

Bacteria of the phylum *Chloroflexi* are morphologically and metabolically highly diverse. To date, very little is known about the secondary metabolism of these microorganisms, providing incentive for a systematic, genomics-driven analysis. In general, the biosynthetic capabilities of *Chloroflexi* bacteria turned out to be rather modest. Only the genera *Herpetosiphon*, *Ktedonobacter* and *Thermogemmatispora* were exceptions to this rule. Several biosynthetic gene clusters from the three genera could not be associated with known natural products, suggesting that they are involved in the production of novel secondary metabolites. The compounds that can be expected from these organisms will be discussed.

Introduction

The phylum *Chloroflexi* was first described by Garrity and Holt in 2001 [1]. It was long considered as an ancient phototrophic lineage. However, recent evidence suggests that the acquisition of photosynthesis genes occurred relatively late in Earth's history [2]. Nowadays, the phylum *Chloroflexi* harbors mainly non-phototrophic organisms, which were reported from soil, decaying organic matter, marine, and freshwater environments [1]. Furthermore, some representatives are commonly found in wastewater treatment plants and metal-contaminated soil [3, 4]. Since the discovery of *Chloroflexus aurantiacus* J-10-fl in 1974 [5], hundreds of *Chloroflexi* and *Chloroflexi*-like strains have been isolated. Analysis of 16S rRNA gene sequences revealed that the phylum *Chloroflexi* is divided into eight classes, namely *Chloroflexia* [1, 6], *Anaerolineae* [7], *Caldilineae* [7], *Ktedonobacteria* [8, 9], *Dehalococcoidetdia* [10], *Thermomicrobia* [11], *Ardenticatenia* [12], and *Thermoflexia* [13]. Each lineage contains only few cultured representatives, but these organisms display a wide range of physiological and morphological properties.

The class *Chloroflexia* comprises the three phototrophic genera *Chloroflexus* [5, 14, 15], *Roseiflexus* [16], and *Oscillochloris* [17, 18], as well as the two non-phototrophic genera *Herpetosiphon* and *Kallotenue*. All species in this class are Gram-negative and multicellular filamentous bacteria with gliding motility. The class *Anaerolineae* contains the largest number of classified species in the phylum *Chloroflexi* [12]. Unlike *Chloroflexia*, members of the class *Anaerolineas* are non-motile bacteria. The class *Caldilinea* consists of aerobic and anaerobic thermophilic bacteria [19, 20]. Species in the class *Ktedonobacteria* exhibit a unique morphology, which is distinct from other bacteria of this phylum. They are characterized by the formation of aerial mycelia and spores [8, 21]. All *Ktedonobacteria* species were assigned as CO-oxidizing bacteria since CO oxidation genes are present in their genomes. The class *Dehalococcoidetdia* consists of organohalide-respiring bacteria, which utilize chlorinated and brominated compounds as respiratory electron acceptors [10, 22-24]. For this reason, *Dehalococcoides*-like strains can be biotechnologically used in order to degrade

chlorinated solvents [25, 26]. The lithoautotrophic nitrite-oxidizing bacterium *Nitrolancea hollandica* was isolated from a nitrifying bioreactor and classified into the class *Thermomicrobia* [27]. With its unique cell wall composition and phylogenetic position, the strain is very different from known nitrite-oxidizing bacteria [28, 29]. In recent years, the two new classes *Ardenticatenia* and *Thermoflexia* have been placed in the phylum *Chloroflexi*. The *Thermoflexia* are distinguished by their high GC content (69.3 mol%), while members of *Ardenticatena* are capable of reducing iron and nitrate [12, 13].

1. Bacterial natural product biosynthesis

Several medicinal drugs have been inspired by bacterial natural products. In particular, Actinobacteria, Proteobacteria and Cyanobacteria served as valuable resources for drug discovery [30-32]. Biosynthetically, natural products can be classified as polyketides, nonribosomal peptides, ribosomally synthesized peptides, terpenoids or hybrids of the aforementioned natural product classes [33]. The former two classes are biosynthesized by nonribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs), respectively.

1.1 Polyketide biosynthesis

PKSs are mechanistically related to fatty acid synthases [36]. They are multi-functional enzymes, which catalyze successive decarboxylative Claisen condensations of acyl-thioester units such as malonyl-CoA and methylmalonyl-CoA. The discovery of the erythromycin biosynthesis gene cluster in 1990 [34] provided important insights into the logic of polyketide biosynthesis. Erythromycin is a 14-membered macrolide, which is produced by three type I PKSs. These enzymes exhibit a modular organization. Each module comprises a set of catalytic domains that are responsible for chain elongation. A minimal PKSs module contains at least three domains: a ketosynthase (KS) domain that catalyzes the decarboxylative Claisen condensation between an extender unit and the growing polyketide chain; an acyl carrier protein (ACP) domain employs a 4'-

phosphopantetheine moiety for covalent binding of the extender unit; an acyltransferase (AT) domain serves as a gatekeeper domain that selects the building block for each chain elongation step and transfers it to the neighboring ACP domain. In addition, some optional domains are involved in the reduction of the β-ketoacyl intermediates. A ketoreductase (KR) domain catalyzes the formation of a secondary alcohol function, which can subsequently be dehydrated upon action of a dehydratase (DH) domain, yielding a double bond. A fully saturated moiety is obtained through the action of an enoylreductase (ER) domain. At the end of a series of reactions, a thioesterase (TE) domain is responsible for the hydrolytic release of the fully processed polyketide chain (Figure 1A).

Several aromatic polyketides, such as tetracyclines, anthracyclines, angucyclines, and benzoisochromanequinones are assembled by type II PKSs [35]. Type II PKS systems consist of ketosynthase units (KS_α and KS_β) and an ACP domain that iteratively catalyze multiple reactions [36]. Type III PKSs belongs to the chalcone/stilbene synthase family. Unlike type I and type II PKSs, type III PKSs catalyze the condensation reaction of acyl-CoA substrates without utilizing an ACP domain [37, 38].

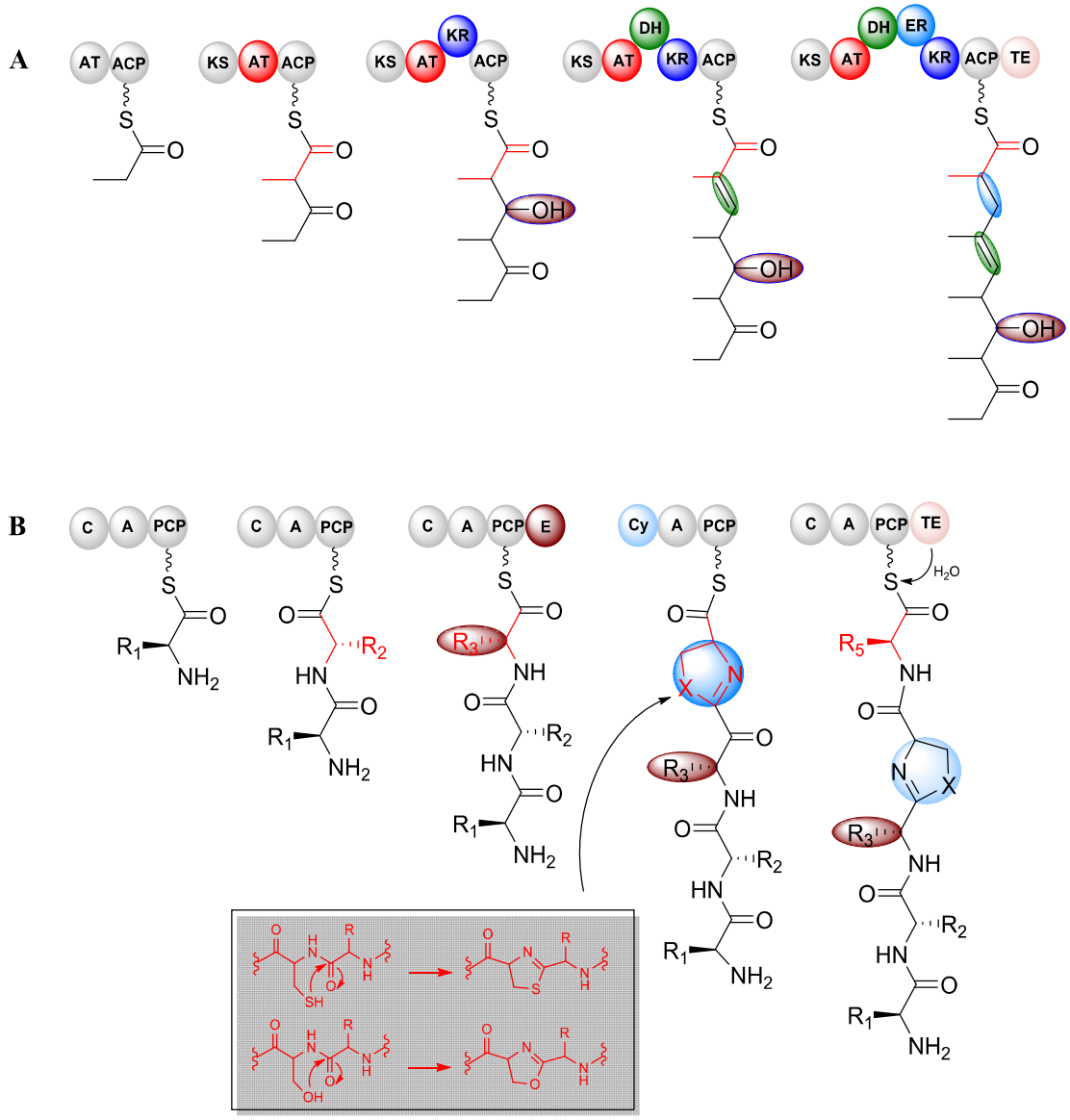


Figure 1. A) Chain extension in a type I PKS assembly line using methylmalonyl-CoA as extender unit; B) Generation of a linear polypeptide by NRPS enzymes.

1.2 Biosynthesis of peptidic natural products

Bacterial nonribosomal peptides include polymyxin, vancomycin, lincomycin and daptomycin [39]. These peptides are produced by multi-enzyme complexes via a thiotemplate mechanism involving NRPSs. The minimal module of an NRPS consists of three catalytic domains, including an adenylation (A), a peptidyl carrier protein (PCP) or thiolation (T) and a condensation (C) domain. The A domain initiates the extension reaction by selecting and activating an amino acid monomer. The thioesterification of the resulting aminoacyl adenylate from the first step is carried out on a

phosphopantetheinylated active site residue of the PCP domain. The C domain is responsible for the formation of a peptide bond [40, 41]. Additional domains with special functions can be found in some modules (Figure 1B). For example, epimerization (E) domains are found in the vancomycin and nocardicin A biosynthesis enzymes [42, 43]. They catalyze L- to D- epimerization. Heterocyclization (Cy) domains catalyze the condensation and intramolecular cyclodehydration of serine, threonine and cysteine residues. A Cy domain is responsible for the formation of a thiazoline ring from cysteine in the biosynthesis of bacillamide E [47]. Similar to PKSs, NRPSs possess thioesterase (TE) domains for the cyclization and release of the final compound [44].

Other peptidic natural products are ribosomally synthesized and posttranslationally modified peptides (RiPPs). They usually contain 20-100 amino acids. All RiPPs are generated from a precursor peptide, which consists of an N-terminal leader and a C-terminal core peptide region. The latter is subject to different types of post-translational modifications (PTM), such as glycosylation, phosphorylation, methylation, cyclization and acetylation. The leader peptide is removed by peptidases or cyclized by PTM to give the final natural product [45]. RiPPs have been subdivided into different classes. Of all classes, lantipeptides, lasso peptides, cyanobactins, and thiopeptides are well known in bacteria. Lantipeptides possess thioether cross-linked rings formed by lanthionine (Lan) and methyllanthionine (MeLan). The unique lariat structure of lasso peptides originates from a linear C-terminal tail and an N-terminal macrolactam ring [46]. Cyanobactins are cyclic RiPPs that are generated by proteolytic cleavage and N to C macrocylization of core peptides [47]. In general, the structure of thiopeptides contains at least one macrocyclic core that includes various modified amino acids, such as dehydroamino acids, oxazoles, thiazoles and thiazolines. The formation of these moieties involves dehydration, dehydrosulfanylation, oxidation and cyclization reactions [48].

1.3 Terpenoid biosynthesis

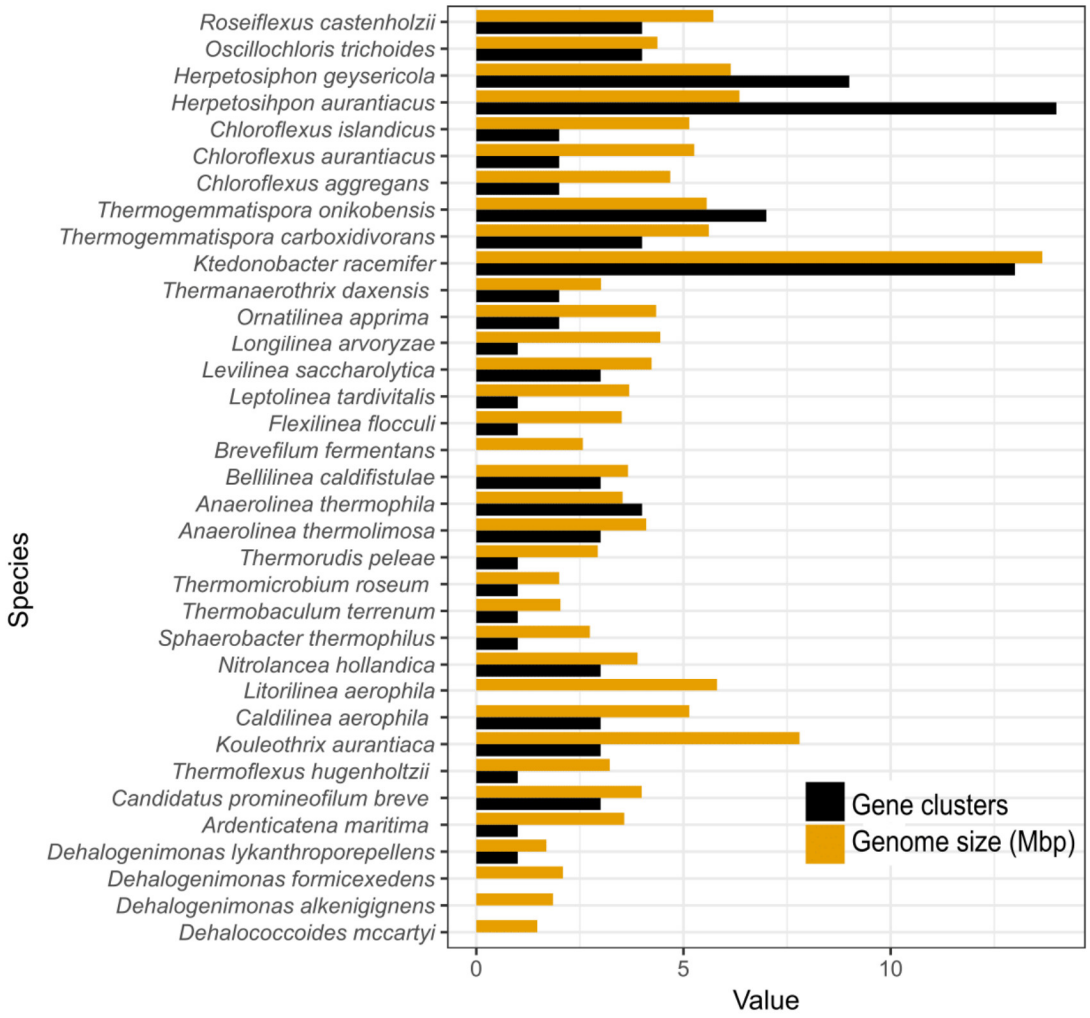
Genes encoding isoprenoid biosynthetic enzymes are widely present in bacteria,

showing their great potential for isoprenoid production. Bacterial isoprenoids display a variety of biological activities, such as antitumor (terpentecin) [49], antibacterial (platensimycin) [50], antioxidant (naphterpin) [51] and anticancer (erythrolic acid) [52]. Terpenoids arise from two precursors, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). The two building blocks can be synthesized via the mevalonate (MEV) pathway or the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway [53]. The structural diversity of terpenoid natural products originates from the combination of different numbers of the aforementioned C_5 isoprenoid precursors. For example, the monoterpenes derive from geranyl diphosphate (GPP), which is formed by the head-to-tail connection of IPP and DMAPP. Farnesyl diphosphate (FPP, C_{15}) is the common precursor of all sesquiterpenes. It is generated by the condensation of GPP with IPP. All diterpenoids derive from geranylgeranyl diphosphate (GGPP, C_{20}), which is generated by an addition of IPP to FPP [54]. Isoprenoids with 25 or 30 carbon atoms are comparatively rare in bacteria, but are commonly found in plants and marine sponges. Carotenoids (C_{40}) are frequently discovered in bacteria, particularly in photosynthetic bacteria. They play a key role in UV protection [55].

2. Biosynthetic potential of the phylum *Chloroflexi*

To identify putative biosynthetic clusters from the phylum *Chloroflexi*, 35 available genome sequences were retrieved from NCBI and analyzed with antiSMASH [56]. The results are depicted in Figure 2. None or few biosynthesis gene clusters were detected in the class *Dehalococcoidetdia*, which might be due to the small genome sizes of the respective bacteria (1.47 to 2.02 Mbp). Similar observations were made for the *Anaerolineae* and *Thermomicrobia*, whose individual genome sizes do not exceed 4.5 Mbp. Every member of these genera features between 0 and 4 biosynthetic loci. The biosynthetically most potent bacteria were found in the classes *Chloroflexia* and *Ktedonobacteria.* In case of the *Chloroflexia*, it was noted that the presence of biosynthesis genes is not evenly distributed among its members. While the

Herpetosiphon genomes are particularly rich in natural product pathways, other genera of this class (*e.g.*, *Chloroflexus*) possess only modest biosynthetic capabilities. With a total of 14 loci, *Herpetosiphon aurantiacus* DSM 785 features even the highest number of biosynthesis gene clusters in the entire phylum. This discovery was surprising, because *H. aurantiacus* DSM 785 has a genome only half the size of *Ktedonobacter racemifer* DSM 44963 (13.66 Mbp). This means that there is no obvious correlation between the genome size and the number of putative biosynthetic gene clusters. According to their genetic endowment *Herpetosiphon* and *Ktedonobacter* are capable to produce all types of natural products, i.e. polyketides, peptides, and terpenoids. *Thermogemmatispora* spp., which are the closest relatives of *Ktedonobacter*, lack PKS gene clusters but possess several NRPS and RiPP gene clusters. In the class *Ktedonobactera*, however, the production of secondary metabolites has been only investigated in *Thermosporothrix hazakensis* to date (Figure 4). The biosynthetic potential of this bacterium is unclear since its genome sequence has not been published. PKS and NRPS gene clusters are absent in the class *Thermomicrobia*, whereas lasso peptide and lantipeptide gene clusters are present in *Thermobaculum terrenum* and *Nitrolancea hollandica*. In addition, the genomes of species in the three genera *Caldilineae*, *Ardenticatena*, and *Thermoflexia* only contain terpene biosynthesis genes. In contrast to the ubiquity of terpene synthase genes, PKS, NRPS, and RiPP biosynthetic pathways are restricted to few *Chloroflexi* bacteria, such as the genera *Herpetosiphon* and *Ktedonobacter*.



204

205 **Figure 2.** Genome sizes and number of secondary metabolite gene clusters in bacteria of the phylum
206 *Chloroflexi*.

207

208 To obtain more information on the terpenoids that can be expected, 42 terpene
209 synthase sequences from cultured and unclassified members of the *Chloroflexi*, and
210 25 presumptive terpene synthases from other phyla were subjected to a phylogenetic
211 analysis. For this, a phylogenetic tree was constructed utilizing MEGA 7.0 [57] with
212 maximum likelihood (ML) estimation (Figure 3). Although no terpene synthases were
213 detected in the five sequenced *Dehalococcoidia* species, such enzymes are present in
214 unclassified members of this class. Phytoene synthases, which are involved in
215 carotenoid biosynthesis, were detected in every class of the phylum *Chloroflexi*.
216 Putative sesquiterpene and diterpene synthases not only occur in the genera

Herpetosiphon, *Roseiflexus*, *Ktedonobacter*, *Longilinea*, and *Nitrolancea*, but were also identified from three unclassified *Chloroflexi* bacteria. The class I terpene synthase lineage comprises a total of ten predicted terpene synthases, including five from *Herpetosiphon*, four from *Roseiflexus* and one from *Ktedonobacter*. Of these terpene synthases, Haur_2987 is a known diterpene synthase, while three other synthases (Haur_2988, RoseRS_3509 and Rcas_0622) are likely responsible for the cyclization of sesquiterpenes [58]. The functions of the putative terpene synthases Haur_4149 and SE18_12445 from *H. aurantiacus* and *H. geysericola* are unclear. The class II contains the known diterpene synthase Haur_2145 from *Herpetosiphon aurantiacus* DSM 785 [59] and six further proteins from different species of the phylum *Chloroflexi*.

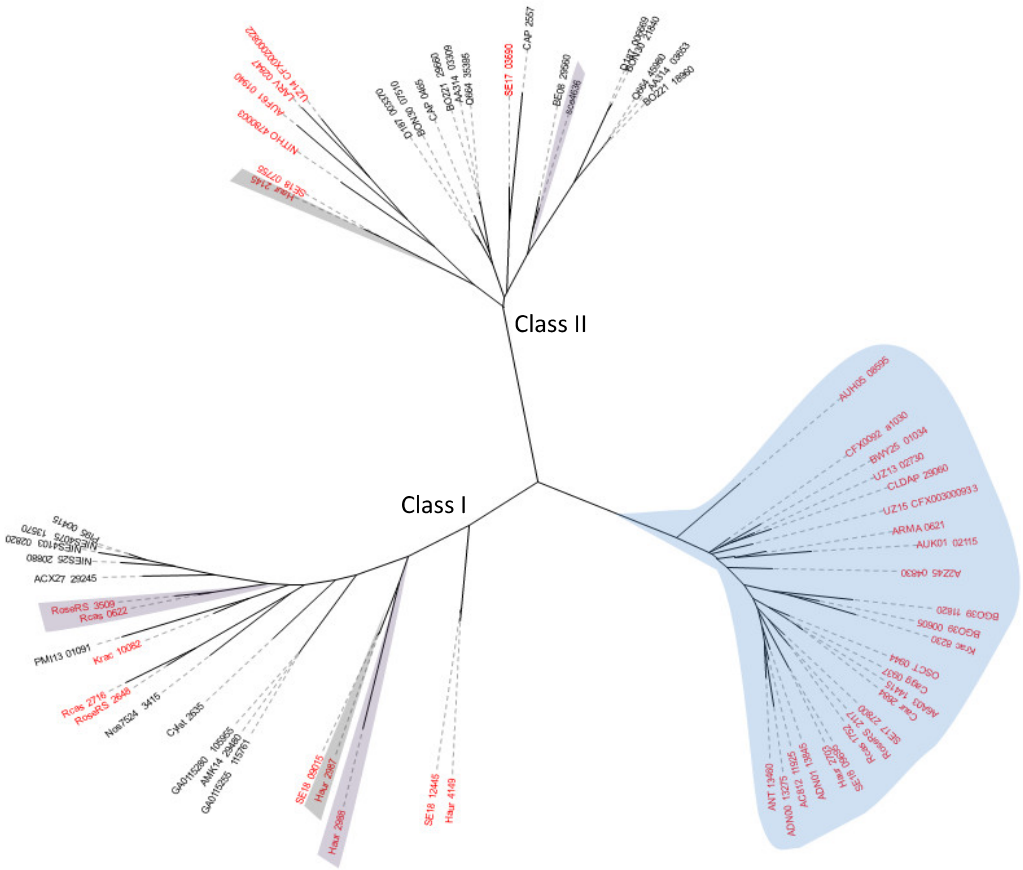


Figure 3. Maximum-likelihood phylogenetic analysis of terpene synthases found in the phylum *Chloroflexi*. Terpene synthases from the phylum *Chloroflexi* are indicated in red-color charters. Phytoene, sesquiterpene and diterpene synthases are highlighted in blue, purple, and grey, respectively.

The metabolic origin of the terpenoid building blocks in the different *Chloroflexi* bacteria was also analyzed. While most bacteria utilize the MEP pathway for IPP and

DMAPP formation [36], the *Chloroflexia*, *Anaerolineae*, *Caldilineae*, *Thermoflexia*, and *Ardenticatenia* were found to exclusively harbor the genes of the MEV pathway. Only the *Thermomicrobia*, *Ktedonobacteria*, and *Dehalococcoidia* were found to possess the MEP pathway.

Although a number of terpene synthases are encoded in the genomes of *Chloroflexi* bacteria, only three isoprenoids have been reported from this phylum to date. Verrucosan-2β-ol, which is also known from marine sponges and liverworts, was the first representative reported from a *Chloroflexus aurantiacus* strain [60]. *O*-methylkolavelool and herpetopanone were recently identified in culture broths of *H. aurantiacus* (Figure 4) [59, 61].

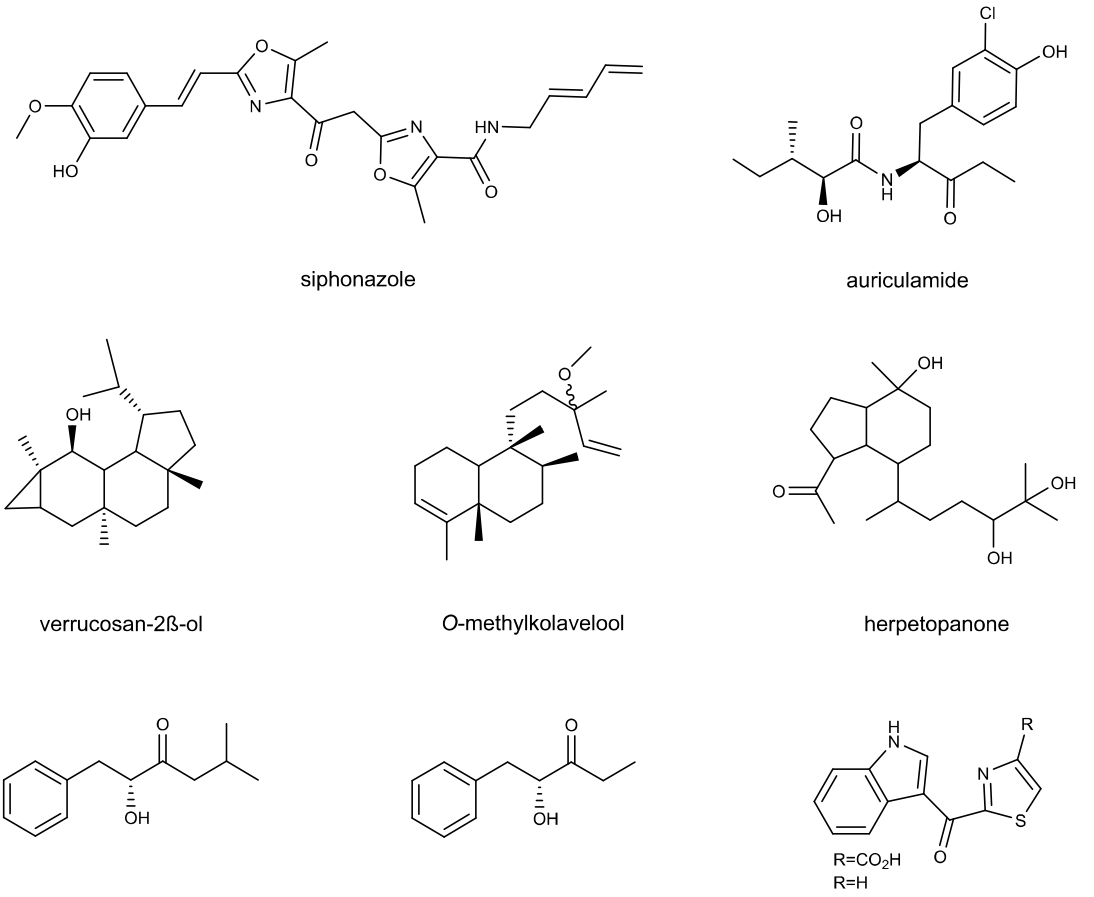


Figure 4. Natural products from the phylum *Chloroflexi*: siphonazole[62], auriculamide[63], *O*-methylkolavelool[59] and herpetopanone[64] were found in *Herpetosiphon* spp.; verrucosan-2β-ol was isolated from *Chloroflexus aurantiacus* J-10-fl[60]; sattabacin, hazakacin[65] and two thiazoles[66] were reported from *Thermosporothrix hazakensis* SK20-1.

2.1 The genus *Herpetosiphon*

The genus *Herpetosiphon* contains four species, *i.e.*, *H. aurantiacus*, *H. geysericola*, *H. giganteus* and *H. gulosus* [61]. However, only the complete genome sequences of *H. aurantiacus* DSM 785 and *H. geysericola* DSM 7119 are currently available. Application of bioinformatics tools including antiSMASH, NRPSpredictor2 and Bagel3 revealed two PKS, four NRPS, three hybrid PKS/NRPS, a thiopeptide, a sactipeptide, and a lantipeptide locus in *H. aurantiacus*. *H. geysericola* possesses two PKS, two NRPS, a hybrid PKS/NRPS, a thiopeptide, and a bacteriocin cluster. A detailed bioinformatics analysis of the annotated loci showed that the two sequenced *Herpetosiphon* species possess the genes for myxochelin biosynthesis (Figure 5).

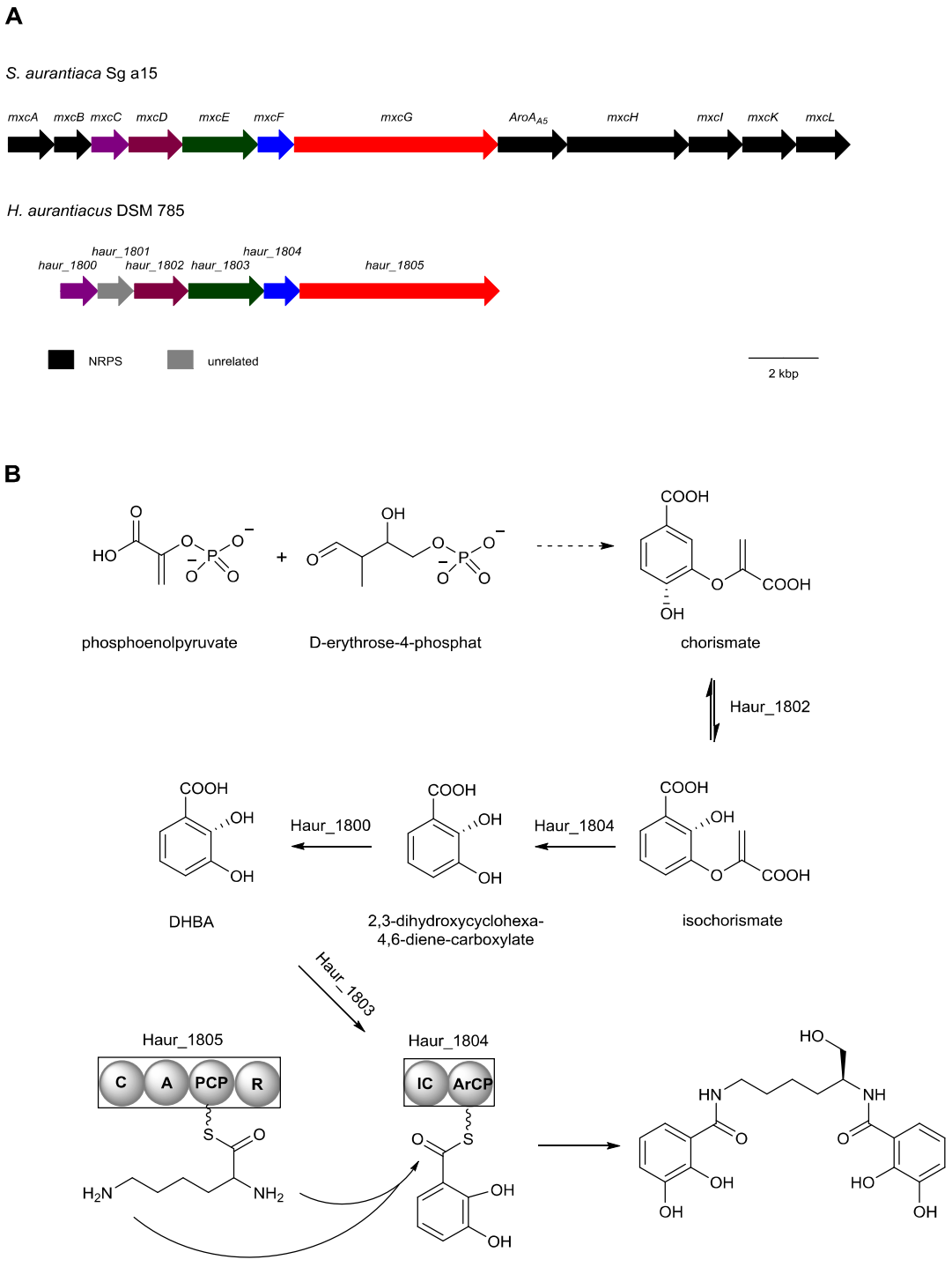


Figure 5. A) The myxochelin gene clusters in *Stigmatella aurantiaca* Sg a15 and *H. aurantiacus* DSM 785; B) Proposed biosynthesis of myxochelin A in *H. aurantiacus*.

This family of siderophore natural products was previously reported from myxobacteria and actinomycetes [67-70]. A direct comparison of the myxochelin loci revealed that the *Herpetosiphon* strains lack the gene for the aminotransferase MxcL,

271 which is involved in the conversion of myxochelin A to myxochelin B. It was hence
272 evident that these organisms are only capable to produce the former metabolite.

273 Auriculamide (Figure 4) is the first natural product isolated from *H. aurantiacus* which
274 is produced via a hybrid PKS/NRPS pathway [69]. The proposed assembly line for the
275 biosynthesis of auriculamide consists of two NRPSs (Haur_2414 and Haur_2415) and
276 a PKSs (Haur_2416). The two NRPSs contain adenylation (A) domains for 2-hydroxy-3-
277 methylvaleric acid (HMVA) and tyrosine substrates. The incorporation of a
278 methylmalonyl-CoA into the peptide chain is catalyzed by Haur_2416. AntiSMASH
279 analysis suggests that the auriculamide biosynthesis locus is larger than previously
280 proposed. Two additional genes encode PKS and NRPS proteins (Haur_2412 and
281 Haur_2413). Hence, the complete locus might be involved in the biosynthesis of an
282 even larger natural product than auriculamide (Figure 6).

283

284 In addition, *H. aurantiacus* possesses pathways to specific biosynthetic building blocks,
285 such as 4-hydroxyphenylglycine (HPG), which is a nonproteinogenic amino acid
286 constituent of glycopeptide antibiotics [71]. The HPG biosynthesis gene cluster is
287 located within a PKS-NRPS gene locus, which consists of one PKS module and eleven
288 NRPS modules. The latter were predicted to activate the amino acids: 4-
289 hydroxyphenylglycin, asparagine, aspartate, glutamic acid, valine, serine, asparagine
290 and threonine (Figure 7). However, the condensation (C) domains in three NRPS
291 modules are likely to be inactive due to the lack of the conserved catalytic motif
292 (HHxxxDG) [72]. The PKS gene (*haur_1873*) shows 40% and 36% similarities to PKS
293 genes involved in puwainaphycin (*puwB*) [73] and in curacin A (*curJ*) [74]
294 biosynthesis.

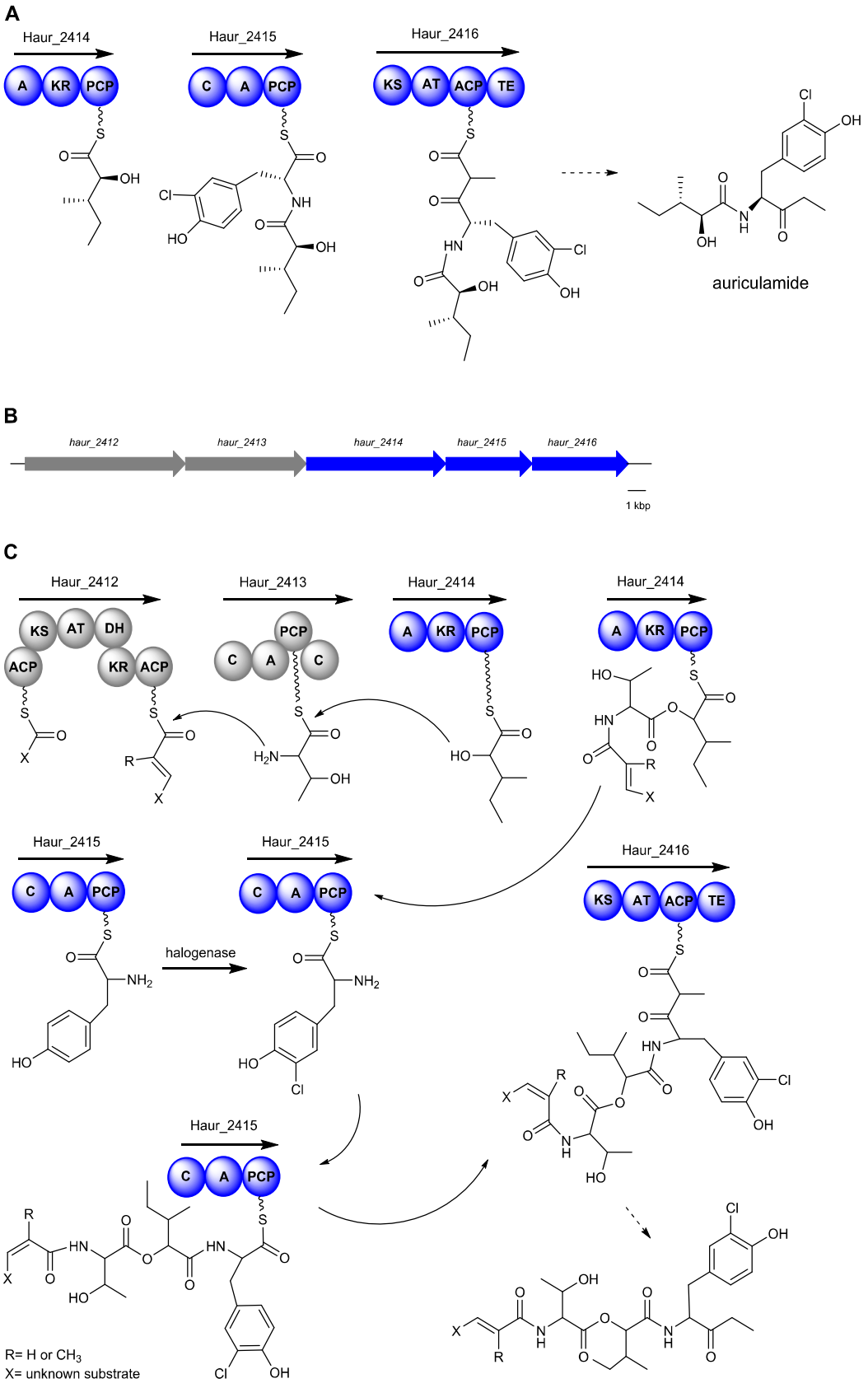
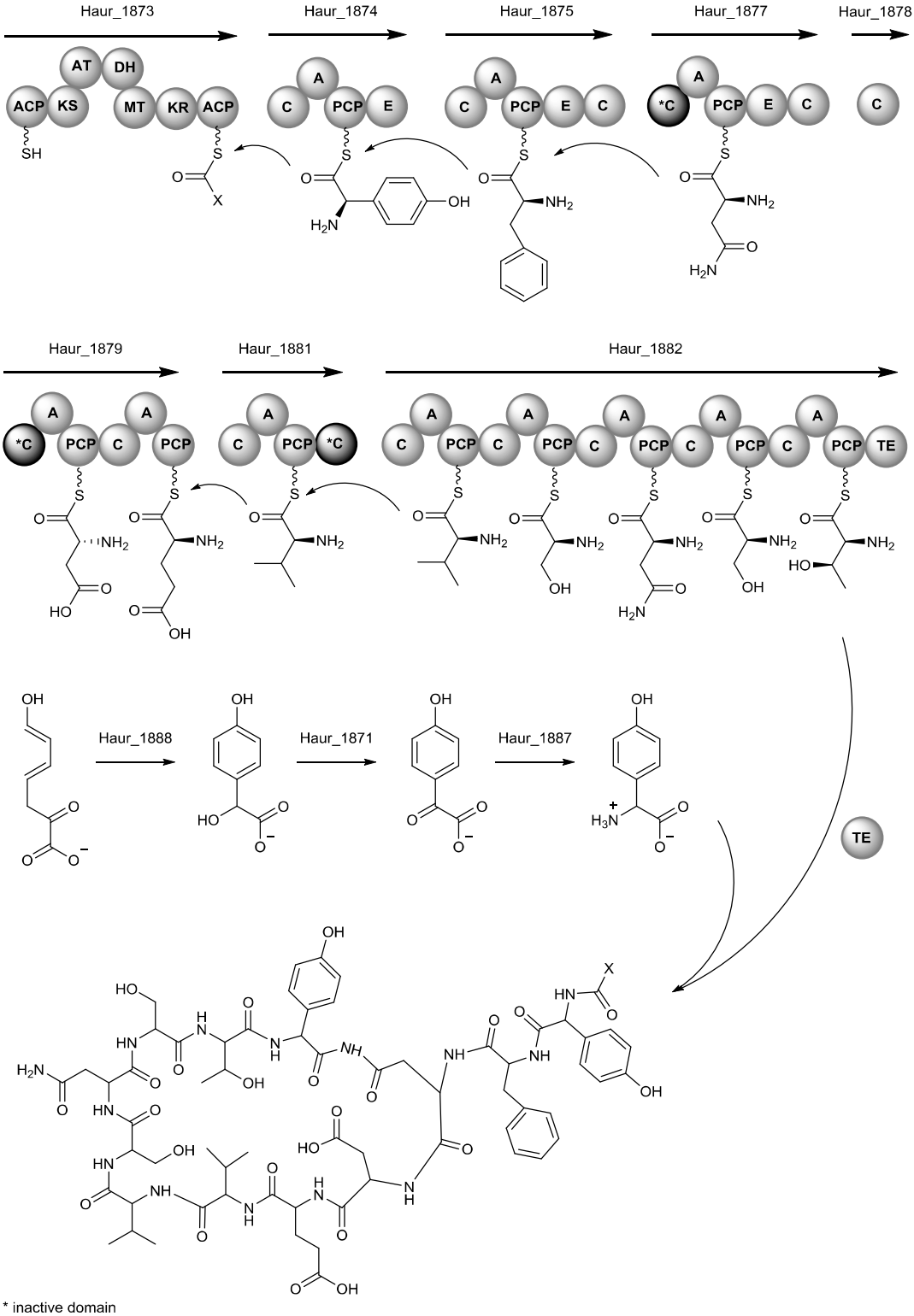


Figure 6. A) Proposed assembly-line biosynthesis of auriculamide; B) The extended gene cluster and C) the hypothetical assembly-line biosynthesis.

299



300

301 **Figure 7**. Proposed assembly-line biosynthesis of an unknown cyclopeptide.

302 Each *Herpetosiphon* species harbors a type III PKS gene cluster, which shows 96%

303 sequence similarity with each other. Furthermore, they show close phylogenetic

relatedness with a type III PKSs from *Bacillus subtilis* (NP_390087), which was previously identified as an alkylpyrone synthase [75]. In view of this, the type III PKSs in the two *Herpetosiphon* strains were proposed to encode the biosynthesis of alkylpyrones (Figure 8).

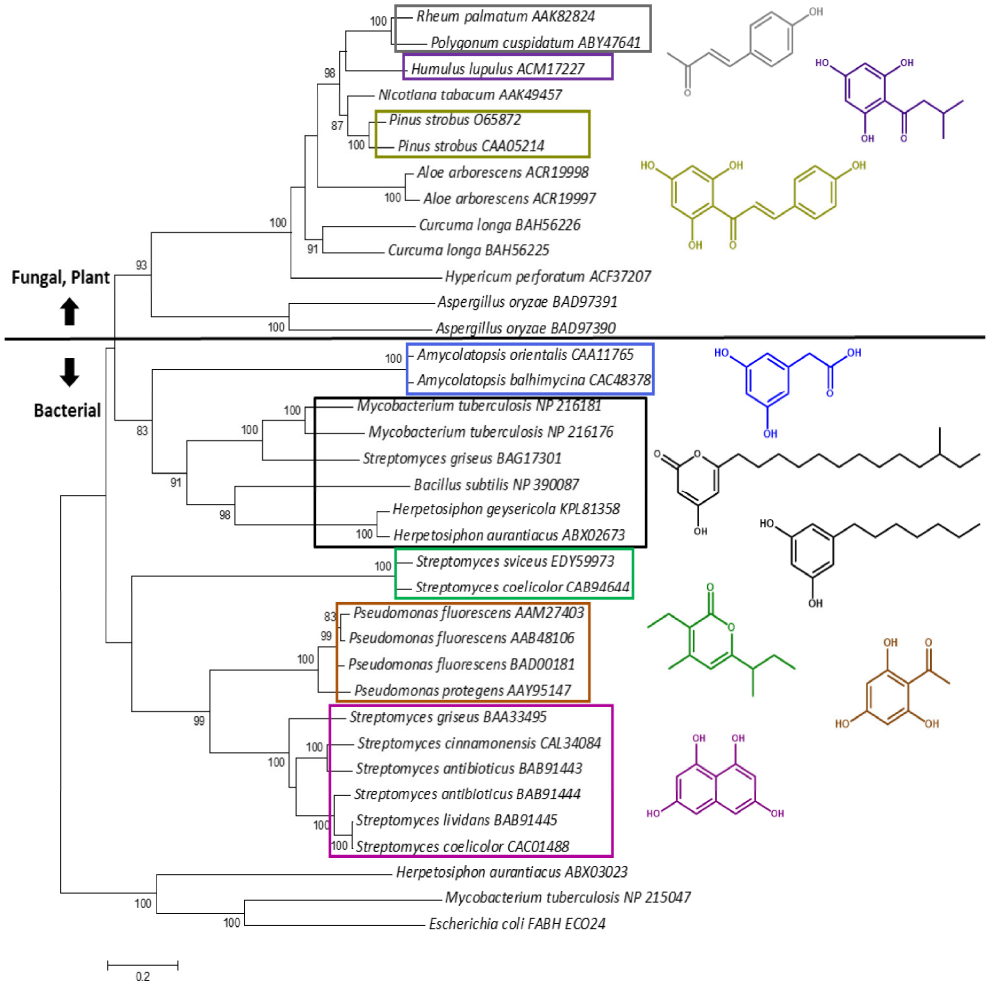


Figure 8. Phylogenetic analysis of bacterial, fungal and plant type III PKSs. The phylogenetic tree was constructed in MEGA 7.0 with the neighbor joining method (1000 bootstrap replicates), based on alignments of the amino acid sequences. Proteins with high similarity were marked with a square and their corresponding metabolic products were used the same color with square.

Furthermore, the genome analysis revealed a large PKS-NRPS gene cluster spanning over 100 kb. This cluster consists of 16 PKS and 7 NRPS modules, which may contribute to the biosynthesis of a lipopeptide (Figure 9). The protein Haur_3971 from *H. aurantiacus* shows homology to the FkbH-like protein PelK [76], which catalyzes the

loading of phosphoglycerate. The three proteins Haur_3968, Haur_3969 and Haur_3970 exhibit high similarity to proteins responsible for the synthesis of glycolate precursors in the pellasoren [76] and soraphen [77] pathways. Haur_3972 and Haur_3973 encode respectively one and six NRPS modules which are predicted to activate ornithine, leucine, alanine, serine and asparagine. Since the genes in this cluster do no show significant similarities to known biosynthesis genes, this cluster likely codes for a new molecule or known metabolite, whose biosynthetic pathway has not been described yet.

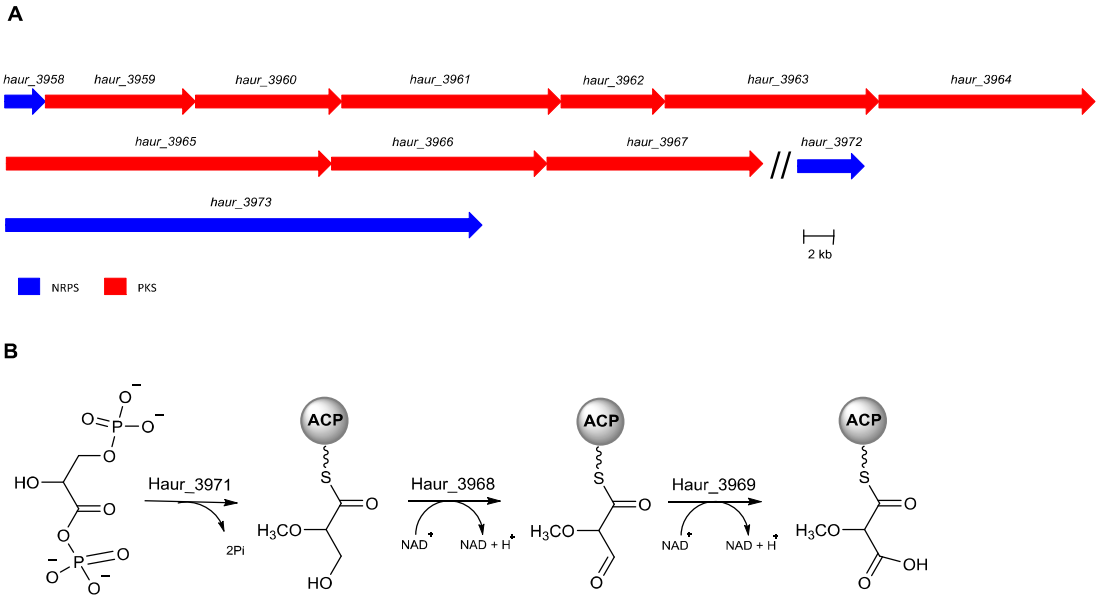


Figure 9. Putative gene cluster (A) and proposed biosynthetic route to a glycolate extender unit (B).

2.2 The genus *Ktedonobacter*

At the time of writing, this genus only contains one species with validly published name. The genome of *K. racemifer* DSM 44963 is the largest in the phylum *Chloroflexi*, encoding 11, 453 putative proteins [78]. As described before, *K. racemifer* possesses a morphology reminiscent of actinobacteria, which are known to produce a variety of bioactive compounds. The ability of *K. racemifer* to produce natural products is unclear since no secondary metabolites have been discovered from this species to date. However, a genome analysis revealed the presence of 13 biosynthesis gene clusters, including 5 PKS, 2 NRPS, 1 hybrid PKS/NRPS, 2 terpenes, 1 lasso peptide, 1

homoserine lactone, and 1 lantipeptide.

A putative 5.97-kb resistomycin-like gene cluster was identified based on homology to biosynthesis enzymes from *Streptomyces resistomycificus*. Resistomycin is a unique pentacyclic polyketide with antibacterial and anti-HIV-1 activities [79]. Its biosynthesis involves a type II PKS, as demonstrated in a heterologous expression experiment. The resistomycin (*rem*) gene cluster in *S. resistomycificus* comprises 18 genes, which are responsible for the biosynthesis, transport, and regulation. Of the biosynthesis proteins, RemA and RemB catalyze the formation of the polyketide chain from malonyl-CoA units. RemF, RemI, and RemL are responsible for cyclization reactions, and RemG and RemH act as methyltransferases [79, 80]. The *rem* operon, which was found in *K. racemifer* DSM 44963, consists of genes *krac_2391* and *krac_2392* (KS_α/KS_β heterodimer), *krac_2388* and *krac_2395* (cyclase), and *krac_2390* (methyltransferase). These genes show strong homology to their characterized counterparts in *S. resistomycificus* (Figure 10). Thus, it is proposed that a resistomycin-like molecule can be produced by this strain.

Another type II PKS gene cluster spanning 3.79 kb consists of four core biosynthesis genes (*krac_1515*, *krac_1516*, *krac_1518* and *krac_1519*). The metabolic product is unclear due to the lack of cyclase and methyltransferase genes. The other biosynthesis loci of *K. racemifer* DSM 44963 lack characterized homologs.

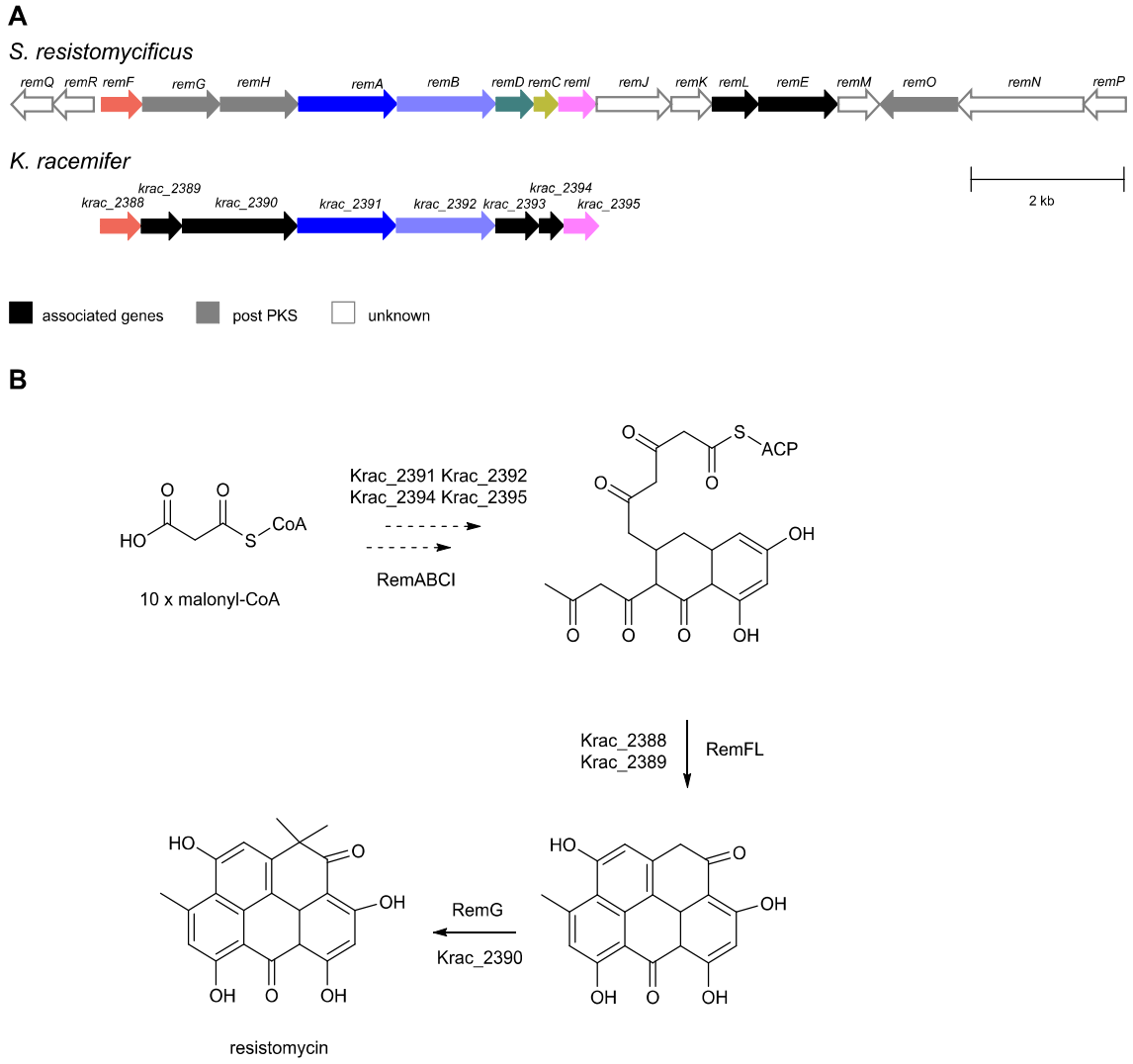


Figure 10. A) Resistomycin biosynthetic gene clusters from *S. resistomycificus* and *K. racemifer*; B) Proposed biosynthesis of resistomycin from *S. resistomycificus* and *K. racemifer*. Homologous genes are shown in the same color. Genes are not homologous of characterized genes but with similar function are marked in grey color.

2.3 The genus *Thermogemmatispora*

The genus *Thermogemmatispora* contains three species: *T. onikobensis*, *T. carboxidivorans*, and *T. foliorum*. The genome sequences of *T. onikobensis* NBRC 111776 and *T. carboxidivorans* DSM 45816 have been released to date. Unlike *K. racemifer*, *T. onikobensis* and *T. carboxidivorans* possess smaller genomes encoding 7 and 4 biosynthetic gene clusters, respectively. *T. onikobensis* has three NRPS, a thiopeptide, a lantipeptide, a lasso peptide, and a phytoene synthase gene cluster, whereas one NRPS, one lantipeptide, one thiopeptide, and one phytoene synthase

cluster were identified from the genome of *T. carboxidivorans*. The lantipeptide and terpene gene clusters from *T. carboxidivorans* and *T. onikobensis* are highly conserved. Most of NRPS gene clusters in these two strains are rather small with only one or two NRPS modules. Only one NRPS gene cluster in *T. carboxidivorans* possesses four modules. The A domains in these modules were predicted to activate leucine, two histidines and valine, respectively. A cytochrome P450 monooxygenase, which is also encoded in this NRPS gene cluster, might be involved in the hydroxylation of the leucine moiety due to its homology to P450 gene (*sce0675*) (45% identity) from *Sorangium cellulosum* So ce56 (Figure 11) [81, 82].

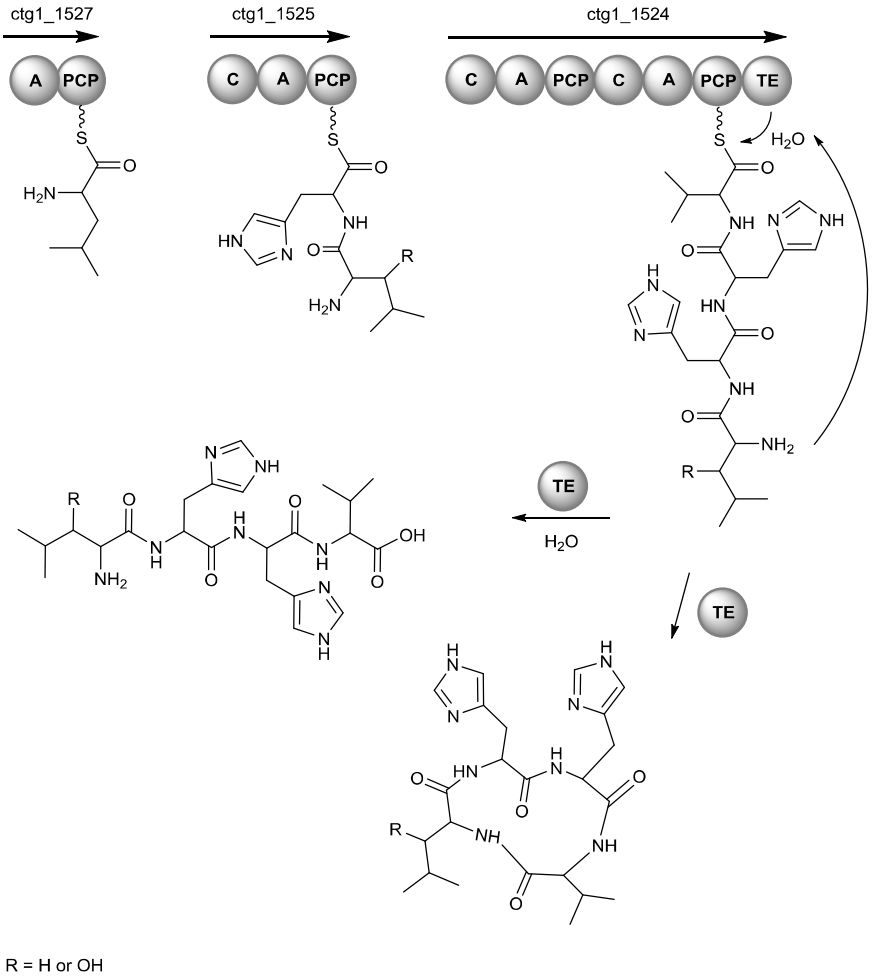


Figure 11. Proposed biosynthetic route to a nonribosomal peptide in *T. carboxidivorans*.

Conclusion

Our analyses revealed that the *Chloroflexi* bacteria are in general capable of natural product biosynthesis. Genome analysis of *K. racemifer* and *Herpetosiphon* species revealed an unexpected potential for the synthesis of novel secondary metabolites. A high number of PKS and NRPS gene clusters were identified from the genera *Herpetosiphon* and *Ktedonobacter*. In addition, genes associated with the biosynthesis of isoprenoids are widely distributed in the phylum *Chloroflexi*, although no monoterpene synthase was detected in this phylum. Phytoene synthases involved in carotenoid biosynthesis are present in every class, except for *Dehalococcoidetdia*. Diterpene and sesquiterpene synthases were mainly discovered in the two genera *Herpetosiphon* and *Roseiflexus*.

Acknowledgements

X. Pan thanks the Chinese Scholarship Council (CSC) for a doctoral stipend.

Reference

1. Garrity GM, Holt JG. Phylum BVI. Chloroflexi phy. nov. In Bergey's Manual of Systematic Bacteriology, 2nd edn, vol. 1, The Archaea and the Deeply Branching and Phototrophic Bacteria, pp. 427– 446. Edited by D. R. Boone, R. W. Castenholz & G. M. Garrity. New York: Springer. **2001**.

2. Shih PM, Ward LM, Fischer WW. Evolution of the 3-hydroxypropionate bicycle and recent transfer of anoxygenic photosynthesis into the *Chloroflexi*. *Proc. Natl. Acad. Sci. U. S. A*.. **2017**;114(40):10749-54.

3. Björnsson L, Hugenholtz P, Tyson GW, Blackall LL. Filamentous *Chloroflexi* (green non-sulfur bacteria) are abundant in wastewater treatment processes with biological nutrient removalc. *Microbiology.* **2002**;148(8):2309-18.

4. Yamada T, Sekiguchi Y, Imachi H, Kamagata Y, Ohashi A, Harada H. Diversity, Localization, and Physiological Properties of Filamentous Microbes Belonging to Chloroflexi Subphylum I in Mesophilic and Thermophilic Methanogenic Sludge Granules. *Appl Environ Microbiol*. **2005**;71(11):7493-503.

5. Pierson BK, Castenholz RW. A phototrophic gliding filamentous bacterium of hot springs, *Chloroflexus aurantiacus*, gen. and sp. nov. *Arch Microbiol.* **1974**;100(1):5-24.

6. Gupta RS, Chander P, George S. Phylogenetic framework and molecular signatures for the class *Chloroflexi* and its different clades; proposal for division of the class *Chloroflexi* class. nov. into the suborder *Chloroflexineae* subord. nov., consisting of the emended family *Oscillochloridaceae* and the family *Chloroflexaceae* fam. nov., and the suborder *Roseiflexineae* subord. nov., containing the family *Roseiflexaceae* fam. nov. Antonie van Leeuwenhoek. **2013**;103(1):99-119.

7. Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H, et al. *Anaerolinea thermolimosa* sp. nov., *Levilinea saccharolytica* gen. nov., sp. nov. and *Leptolinea tardivitalis* gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes *Anaerolineae* classis nov. and *Caldilineae* classis nov. in the bacterial phylum *Chloroflexi*. *Int. J. Syst. Evol. Microbiol*. **2006**;56(6):1331-40.

8. Yabe S, Aiba Y, Sakai Y, Hazaka M, Yokota A. *Thermosporothrix hazakensis* gen. nov., sp. nov.,

isolated from compost, description of *Thermosporotrichaceae* fam. nov. within the class *Ktedonobacteria* Cavaletti et al. 2007 and emended description of the class *Ktedonobacteria*. *Int. J. Syst. Evol. Microbiol*. **2010**;60(8):1794-801.

9. Cavaletti L, Monciardini P, Bamonte R, Schumann P, Rohde M, Sosio M, et al. New lineage of filamentous, spore-forming, gram-positive bacteria from soil. *Appl Environ Microbiol*. **2006**;72(6):4360-9.

10. Löffler FE, Yan J, Ritalahti KM, Adrian L, Edwards EA, Konstantinidis KT, et al. *Dehalococcoides mccartyi* gen. nov., sp. nov., obligately organohalide-respiring anaerobic bacteria relevant to halogen cycling and bioremediation, belong to a novel bacterial class, *Dehalococcoidia* classis nov., order *Dehalococcoidales* ord. nov. and family *Dehalococcoidaceae* fam. nov., within the phylum *Chloroflexi*. *Int. J. Syst. Evol. Microbiol*. **2013**;63(2):625-35.

11. Garrity GM, Holt JG, Perry JJ. Phylum BVII. Thermomicrobia phy. nov. In: Boone DR, Castenholz RW, Garrity GM, editors. Bergey's Manual® of Systematic Bacteriology: Volume One : The Archaea and the Deeply Branching and Phototrophic Bacteria. New York, NY: Springer New York; **2001**. p. 447-50.

12. Kawaichi S, Ito N, Kamikawa R, Sugawara T, Yoshida T, Sako Y. *Ardenticatena maritima* gen. nov., sp. nov., a ferric iron- and nitrate-reducing bacterium of the phylum '*Chloroflexi*' isolated from an iron-rich coastal hydrothermal field, and description of *Ardenticatenia* classis nov. *Int. J. Syst. Evol. Microbiol.* **2013**;63(8):2992-3002.

13. Dodsworth JA, Gevorkian J, Despujos F, Cole JK, Murugapiran SK, Ming H, et al. *Thermoflexus hugenholtzii* gen. nov., sp. nov., a thermophilic, microaerophilic, filamentous bacterium representing a novel class in the *Chloroflexi*, *Thermoflexia* classis nov., and description of *Thermoflexaceae* fam. nov. and *Thermoflexales* ord. nov. *Int. J. Syst. Evol. Microbiol*. **2014**;64(6):2119-27.

14. Hanada S, Hiraishi A, Shimada K, Matsuura K. *Chloroflexus aggregans* sp. nov., a filamentous phototrophic bacterium which forms dense cell aggregates by active gliding movement. *Int. J. Syst. Evo.l* Microbiol. **1995**;45(4):676-81.

15. Gaisin VA, Kalashnikov AM, Grouzdev DS, Sukhacheva MV, Kuznetsov BB, Gorlenko VM. *Chloroflexus islandicus* sp. nov., a thermophilic filamentous anoxygenic phototrophic bacterium from a geyser. *Int. J. Syst Evol. Microbiol*. **2017**;67(5):1381-6.

16. Hanada S, Takaichi S, Matsuura K, Nakamura K. *Roseiflexus castenholzii* gen. nov., sp. nov., a thermophilic, filamentous, photosynthetic bacterium that lacks chlorosomes. *Int. J. Syst. Evol. Microbiol.*

2002;52(Pt 1):187-93.

17. Keppen OI, Baulina OI, Kondratieva EN. Oscillochloris trichoides neotype strain DG-6. *Photosynth Res*. **1994**;41(1):29-33.

18. Keppen OI, Tourova TP, Kuznetsov BB, Ivanovsky RN, Gorlenko VM. Proposal of *Oscillochloridaceae* fam. nov. on the basis of a phylogenetic analysis of the filamentous anoxygenic phototrophic bacteria, and emended description of *Oscillochloris* and *Oscillochloris trichoides* in comparison with further new isolates. *Int. J. Syst. Evol. Microbiol*. **2000**;50(4):1529-37.

19. Grégoire P, Bohli M, Cayol J-L, Joseph M, Guasco S, Dubourg K, et al. *Caldilinea tarbellica* sp. nov., a filamentous, thermophilic, anaerobic bacterium isolated from a deep hot aquifer in the Aquitaine Basin. *Int. J. Syst. Evol. Microbiol*. **2011**;61(6):1436-41.

20. Kale V, Björnsdóttir SH, Friðjónsson ÓH, Pétursdóttir SK, Ómarsdóttir S, Hreggviðsson GÓ. *Litorilinea aerophila* gen. nov., sp. nov., an aerobic member of the class *Caldilineae*, phylum *Chloroflexi*, isolated from an intertidal hot spring. *Int. J. Syst. Evol. Microbiol*. **2013**;63(3):1149-54.

21. Yabe S, Sakai Y, Abe K, Yokota A. Diversity of Ktedonobacteria with Actinomycetes-Like Morphology in Terrestrial Environments. *Microbes Environ*. **2017**;32(1):61-70.

22. Moe WM, Yan J, Nobre MF, da Costa MS, Rainey FA. *Dehalogenimonas lykanthroporepellens* gen. nov., sp. nov., a reductively dehalogenating bacterium isolated from chlorinated solvent-contaminated groundwater. *Int. J. Syst. Evol. Microbiol*. **2009**;59(11):2692-7.

23. Bowman KS, Nobre MF, da Costa MS, Rainey FA, Moe WM. *Dehalogenimonas alkenigignens* sp. nov., a chlorinated-alkane-dehalogenating bacterium isolated from groundwater. *Int. J. Syst. Evol. Microbiol*. **2013**;63(4):1492-8.

24. Key TA, Bowman KS, Lee I, Chun J, Albuquerque L, da Costa MS, et al. *Dehalogenimonas formicexedens* sp. nov., a chlorinated alkane-respiring bacterium isolated from contaminated groundwater. *Int. J. Syst. Evol. Microbiol.* **2017**;67(5):1366-73.

25. Cupples AM, Spormann AM, McCarty PL. Comparative Evaluation of Chloroethene Dechlorination to Ethene by *Dehalococcoides*-like Microorganisms. *Environ. Scl. Technol*. **2004**;38(18):4768-74.

26. Löffler FE, Ritalahti KM, Zinder SH. *Dehalococcoides* and reductive dechlorination of chlorinated solvents. Bioaugmentation for groundwater remediation: Springer; **2013**. p. 39-88.

27. Sorokin DY, Vejmelkova D, Lücker S, Streshinskaya GM, Rijpstra WIC, Sinninghe Damsté JS, et al. *Nitrolancea hollandica* gen. nov., sp. nov., a chemolithoautotrophic nitrite-oxidizing bacterium isolated

518 from a bioreactor belonging to the phylum *Chloroflexi*. *Int. J. Syst. Evol. Microbiol*. **2014**;64(6):1859-65.
519 28. Spieck E, Bock E. The Lithoautotrophic Nitrite-Oxidizing Bacteria. In: Brenner DJ, Krieg NR, Staley
520 JT, Garrity GM, editors. Bergey's Manual® of Systematic Bacteriology: Volume Two: The Proteobacteria,
521 Part A Introductory Essays. Boston, MA: Springer US; **2005**. p. 149-53.
522 29. Luecker S, Nowka B, Rattei T, Spieck E, Daims H. The Genome of *Nitrospina gracilis* Illuminates the
523 Metabolism and Evolution of the Major Marine Nitrite Oxidizer. *Front Microbiol*. **2013**;4(27).
524 30. De Carvalho CC, Fernandes P. Production of metabolites as bacterial responses to the marine
525 environment. *Mar Drugs*. **2010**;8(3):705-27.
526 31. Bérdy J. Thoughts and facts about antibiotics: where we are now and where we are heading. *J.*
527 *Antibiot (Tokyo)*. **2012**;65(8):385-95.
528 32. Vijayakumar S, Menakha M. Pharmaceutical applications of cyanobacteria—A review. *J. Acute*
529 *Medicine*. **2015**;5(1):15-23.
530 33. Dewick PM. Medicinal natural products: a biosynthetic approach: John Wiley & Sons; **2002**.
531 34. Cortes J, Haydock SF, Roberts GA, Bevitt DJ, Leadlay PF. An unusually large multifunctional
532 polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*.
533 *Nature*. **1990**;348(6297):176-8.
534 35. Zhou H, Li Y, Tang Y. Cyclization of aromatic polyketides from bacteria and fungi. *Nat. Prod. Rep*.
535 **2010**;27(6):839-68.
536 36. Zhan J. Biosynthesis of bacterial aromatic polyketides. *Curr. Top. Med. Chem*. **2009**;9(17):1598-610.
537 37. Yu D, Xu F, Zeng J, Zhan J. Type III polyketide synthases in natural product biosynthesis. *IUBMB life*.
538 **2012**;64(4):285-95.
539 38. Shimizu Y, Ogata H, Goto S. Type III polyketide synthases: functional classification and
540 phylogenomics. *ChemBioChem*. **2017**;18(1):50-65.
541 39. Felnagle EA, Jackson EE, Chan YA, Podevels AM, Berti AD, McMahon MD, et al. Nonribosomal
542 peptide synthetases involved in the production of medically relevant natural products. *Mol. Pharm*.
543 **2008**;5(2):191-211.
544 40. Challis GL, Naismith JH. Structural aspects of non-ribosomal peptide biosynthesis. *Curr. Opin. Struct*
545 *Biol*. **2004**;14(6):748-56.
546 41. Bloudoff K, Schmeing TM. Structural and functional aspects of the nonribosomal peptide
547 synthetase condensation domain superfamily: discovery, dissection and diversity. *Biochim. Biophys.*

548 *Acta*. **2017**;1865(11, Part B):1587-604.
549 42. Schmartz PC, Zerbe K, Abou-Hadeed K, Robinson JA. Bis-chlorination of a hexapeptide-PCP
550 conjugate by the halogenase involved in vancomycin biosynthesis. *Org. Biomol. Chem*.
551 **2014**;12(30):5574-7.
552 43. Gunsior M, Breazeale SD, Lind AJ, Ravel J, Janc JW, Townsend CA. The Biosynthetic Gene Cluster
553 for a Monocyclic β-Lactam Antibiotic, Nocardicin A. *Chem. Biol*. **2004**;11(7):927-38.
554 44. Croteau R, Davis E, Hartmann T, Hemscheidt T, Sanz-Cervera J, Shen B, et al. Biosynthesis: Aromatic
555 Polyketides, Isoprenoids, Alkaloids: Springer; **2003**.
556 45. McIntosh JA, Donia MS, Schmidt EW. Ribosomal peptide natural products: bridging the ribosomal
557 and nonribosomal worlds. *Nat. Prod. Rep*. **2009**;26(4):537-59.
558 46. Hegemann JD, Zimmermann M, Xie X, Marahiel MA. Lasso peptides: an intriguing class of bacterial
559 natural products. *Acc. Chem. Res*. **2015**;48(7):1909-19.
560 47. Sivonen K, Leikoski N, Fewer DP, Jokela J. Cyanobactins—ribosomal cyclic peptides produced by
561 cyanobacteria. *Appl. Microbiol. Biotechnol*. **2010**;86(5):1213-25.
562 48. Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, et al. Ribosomally synthesized and
563 post-translationally modified peptide natural products: overview and recommendations for a universal
564 nomenclature. *Nat. Prod. Rep*. **2013**;30(1):108-60.
565 49. Dairi T, Hamano Y, Kuzuyama T, Itoh N, Furihata K, Seto H. Eubacterial diterpene cyclase genes
566 essential for production of the isoprenoid antibiotic terpentecin. *J. Bacteriol.* **2001**;183(20):6085-94.
567 50. Smanski MJ, Peterson RM, Rajski SR, Shen B. Engineered Streptomyces platensis strains that
568 overproduce antibiotics platensimycin and platencin. *Antimicrob. Agents Chemother*. **2009**;53(4):1299-
569 304.
570 51. Shin-ya K, Imai S, Furihata K, Hayakawa Y, Kato Y, Vanduyne GD, et al. Isolation and structural
571 elucidation of an antioxidative agent, naphterpin. *J. Antibiot (Tokyo)*. **1990**;43(4):444-7.
572 52. Hu Y, Legako AG, Espindola APDM, MacMillan JB. Erythrolic acids A–E, Meroterpenoids from a
573 Marine-Derived Erythrobacter sp. *J. Org. Chem*. **2012**;77(7):3401-7.
574 53. Kuzuyama T. Mevalonate and nonmevalonate pathways for the biosynthesis of isoprene units.
575 *Biosc., biotechno., Biochem*. **2002**;66(8):1619-27.
576 54. Dewick PM. The biosynthesis of C 5–C 25 terpenoid compounds. *Nat. Prod. Rep*. **2002**;19(2):181-
577 222.

55. Galasso C, Corinaldesi C, Sansone C. Carotenoids from Marine Organisms: Biological Functions and Industrial Applications. *Antioxidants*. **2017**;6(4):96.

56. Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, et al. antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. *Nucleic. Acid. Res*. **2013**;41(1):204-12.

57. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol*. **2016**;33(7):1870-4.

58. Yamada Y, Kuzuyama T, Komatsu M, Shin-Ya K, Omura S, Cane DE, et al. Terpene synthases are widely distributed in bacteria. *Proc. Natl. Acad. Sci. U. S. A*. **2015**;112(3):857-62.

59. Nakano C, Oshima M, Kurashima N, Hoshino T. Identification of a New Diterpene Biosynthetic Gene Cluster that Produces O-Methylkolavelool in *Herpetosiphon aurantiacus*. *ChemBioChem*. **2015**;16(5):772-81.

60. Hefter J, Richnow HH, Fischer U, Trendel JM, Michaelis W. (-)-Verrucosan-2β-ol from the phototrophic bacterium *Chloroflexus aurantiacus*: first report of a verrucosane-type diterpenoid from a prokaryote. *Microbiology*. **1993**;139(11):2757-61.

61. Pan X, Kage H, Martin K, Nett M. *Herpetosiphon gulosus* sp. nov., a filamentous predatory bacterium isolated from sandy soil and *Herpetosiphon giganteus* sp. nov., nom. rev. *Int. J. Syst. Evol. Microbiol*. **2017**;67(7):2476-81.

62. Nett M, Erol Ö, Kehraus S, Köck M, Krick A, Eguereva E, et al. Siphonazole, an unusual metabolite from *Herpetosiphon* sp. *Angew. Chem. Int. Ed* **2006**;45(23):3863-7.

63. Schieferdecker S, Domin N, Hoffmeier C, Bryant DA, Roth M, Nett M. Structure and Absolute Configuration of Auriculamide, a Natural Product from the Predatory Bacterium *Herpetosiphon aurantiacus*. *Eur. J. Org. Chem*. **2015**; 14: 3057-62.

64. Pan X, Domin N, Schieferdecker S, Kage H, Roth M, Nett M. Herpetopanone, a diterpene from *Herpetosiphon aurantiacus* discovered by isotope labeling. *Beilstein J. Org. Chem*. **2017**;13:2458.

65. Park JS, Kagaya N, Hashimoto J, Izumikawa M, Yabe S, Shin-ya K, et al. Identification and Biosynthesis of New Acyloins from the Thermophilic Bacterium *Thermosporothrix hazakensis* SK20-1T. *ChemBioChem*. **2014**;15(4):527-32.

66. Park J-S, Yabe S, Shin-ya K, Nishiyama M, Kuzuyama T. New 2-(1'*H*-indole-3'-carbonyl)-thiazoles derived from the thermophilic bacterium *Thermosporothrix hazakensis* SK20-1T. *J. Antibiot (Tokyo)*.

608 **2015**;68(1):60.

609 67. Kunze B, Bedorf N, Kohl W, Hofle G, Reichenbach H. Myxochelin A, a new iron-chelating compound
610 from *Angiococcus disciformis* (Myxobacterales). Production, isolation, physico-chemical and biological
611 properties. *J. Antibiot (Tokyo)*. **1989**;42(1):14-7.

612 68. Silakowski B, Kunze B, Nordsiek G, Blocker H, Hofle G, Muller R. The myxochelin iron transport
613 regulon of the myxobacterium *Stigmatella aurantiaca* Sg a15. *Eur. J. Biochem*. **2000**;267(21):6476-85.

614 69. Schieferdecker S, Domin N, Hoffmeier C, Bryant DA, Roth M, Nett M. Structure and Absolute
615 Configuration of Auriculamide, a Natural Product from the Predatory Bacterium *Herpetosiphon*
616 *aurantiacus*. *Eur J. Org. Chem*. **2015**:3057-62.

617 70. Miyanaga S, Obata T, Onaka H, Fujita T, Saito N, Sakurai H, et al. Absolute configuration and
618 antitumor activity of myxochelin A produced by *Nonomuraea pusilla* TP-A0861. *J. Antibiot (Tokyo)*.
619 **2006**;59(11):698.

620 71. Kastner S, Müller S, Natesan L, König GM, Guthke R, Nett M. 4-Hydroxyphenylglycine biosynthesis
621 in *Herpetosiphon aurantiacus*: a case of gene duplication and catalytic divergence. *Arch. Microbiol*.
622 **2012**;194(6):557-66.

623 72. Marahiel MA, Stachelhaus T, Mootz HD. Modular peptide synthetases involved in nonribosomal
624 peptide synthesis. *Chem. Rev*. **1997**;97(7):2651-74.

625 73. Mareš J, Hájek J, Urajová P, Kopecký J, Hrouzek P. A hybrid non-ribosomal peptide/polyketide
626 synthetase containing fatty-acyl ligase (FAAL) synthesizes the β-amino fatty acid lipopeptides
627 puwainaphycins in the Cyanobacterium *Cylindrospermum alatosporum*. *PloS one*. **2014**;9(11):e111904.

628 74. Chang Z, Sitachitta N, Rossi JV, Roberts MA, Flatt PM, Jia J, et al. Biosynthetic Pathway and Gene
629 Cluster Analysis of Curacin A, an Antitubulin Natural Product from the Tropical Marine Cyanobacterium
630 Lyngbya m ajuscula. *J. Nat. Prod*. **2004**;67(8):1356-67.

631 75. Nakano C, Ozawa H, Akanuma G, Funa N, Horinouchi S. Biosynthesis of aliphatic polyketides by
632 type III polyketide synthase and methyltransferase in *Bacillus subtilis*. *J. Bacteriol*. **2009**;191(15):4916-
633 23.

634 76. Jahns C, Hoffmann T, Müller S, Gerth K, Washausen P, Höfle G, et al. Pellasoren: structure
635 elucidation, biosynthesis, and total synthesis of a cytotoxic secondary metabolite from *Sorangium*
636 *cellulosum*. *Angew. Chem. Int. Edi*. **2012**;51(21):5239-43.

637 77. Ligon J, Hill S, Beck J, Zirkle R, Molnár I, Zawodny J, et al. Characterization of the biosynthetic gene

638 cluster for the antifungal polyketide soraphen A from *Sorangium cellulosum* So ce26. *Gene*.
639 **2002**;285(1):257-67.
640 78. Chang Y-J, Land M, Hauser L, Chertkov O, Del Rio TG, Nolan M, et al. Non-contiguous finished
641 genome sequence and contextual data of the filamentous soil bacterium *Ktedonobacter racemifer* type
642 strain (SOSP1-21). *Stand Genomic Sci*. **2011**;5.
643 79. Jakobi K, Hertweck C. A gene cluster encoding resistomycin biosynthesis in Streptomyces
644 resistomycificus; exploring polyketide cyclization beyond linear and angucyclic patterns. *J. Am. Chem.*
645 *Soc*. **2004**;126(8):2298-9.
646 80. Fritzsche K, Ishida K, Hertweck C. Orchestration of discoid polyketide cyclization in the resistomycin
647 pathway. *J. Am. Chem. Soc*. **2008**;130(26):8307-16.
648 81. Kern F, Khatri Y, Litzenburger M, Bernhardt R. CYP267A1 and CYP267B1 from *Sorangium cellulosum*
649 So ce56 are highly versatile drug metabolizers. *Drug Metab. Dispos*. **2016**;44(4):495-504.
650 82. Khatri Y, Hannemann F, Ewen KM, Pistorius D, Perlova O, Kagawa N, et al. The CYPome of *Sorangium*
651 *cellulosum* So ce56 and identification of CYP109D1 as a new fatty acid hydroxylase. *Chem. Biol*.
652 **2010**;17(12):1295-305.
653

4 Additional results

4.1 Identification and classification of *Herpetosiphon* sp.

The *Herpetosiphon* strains DSM 52868 and DSM 52870 were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ). DSM 52868 had previously been isolated from forest soil in Germany, whereas DSM 52870 had been isolated from soil in the USA. Both strains show the same morphological and phenotypical features as the type strains *H. aurantiacus* DSM 785 and *H. geysericola* DSM 7119. Effects of pH, temperature, NaCl concentration, and antibiotics on the growth of *Herpetosiphon* spp. DSM 52868 and DSM 52870 were determined in VY/2 medium. To determine the pH range for growth, the medium was adjusted to pH 4.0-9.0 with HCl and NaOH and incubation was conducted at 30 °C. To evaluate the temperature range for growth, the two strains were incubated at 4, 20, 26, 30, 37 and 45 °C (pH 7.2). The halotolerance was tested at 1, 2, 3, 4, or 5% (w/v) NaCl. Effect of antibiotics on growth was tested by placing antibiotic discs (BBLTM Sensi-Disc™, Becton Dickinson) on VY/2 agar plates for 3 to 4 days at 28 °C. The results of these analyses are shown in Table 2.

Table 2. Phenotypic characteristics of *Herpetosiphon* strains DSM 52868 (1) and DSM 52870 (2) in comparison to the type strain *H. giganteus* DSM 589 (3). +, positive; -, negative

Characteristic	**1**	**2**	**3**
pH for growth	6-9	6-9	6-9
Temperature for growth (°C)	20-37	20-37	20-37
NaCl tolerance (%)	3	3	3
Resistance to:			
nalidixic acid (30 µg)	+	-	-
oxytetracyclin (30 µg)	+	-	-
sulfonamide (200 µg)	+	-	-
gentamicin (10 µg)	-	-	+

To establish the phylogenetic relationships between *Herpetosiphon* spp. and related members of *Chloroflexi*, sequence comparison of 16S rRNA was performed by BLAST. Since both strains show 99% sequence similarity with *Herpetosiphon giganteus* Hp a2, *Herpetosiphon* sp. DSM 52868 and *Herpetosiphon* sp. DSM 52870 were placed into the same lineage. A taxonomic description of *H. giganteus* is included in manuscript B.

4.2 Isolation of natural products

4.2.1 General information

High resolution mass analyses were performed with an Executive Mass Spectrometer (Thermo-Scientific). 1D and 2D NMR measurements were carried out at 300 K on Bruker Advance III 500 and 600 MHz spectrometers. Dimethylsulfoxide (DMSO)-d_6 was used as solvent and internal standard (δ_H = 2.50 ppm; δ_C = 39.5 ppm). Semipreparative HPLC was conducted with a Shimadzu HPLC system (LC-20AT, SPD-M20A).

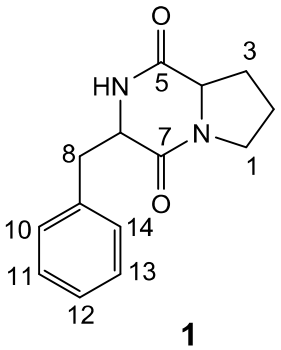
4.2.2 Isolation of diketopiperazines from *H. aurantiacus* DSM 785

H. aurantiacus DSM 785 was grown as a shaking culture (90 rpm) at a 10 L scale in Hp74 broth (L-glutamic acid monosodium salt monohydrate, 1% (w/v); yeast extract, 0.2% (w/v); $MgSO_4$ x 7 H_2O, 0.2% (w/v)) and a temperature of 30 °C. After seven days of cultivation, the culture broth was centrifuged at 7,000 rpm for 15 min. The supernatant was mixed with Amberlite XAD-2 resin (10%, w/v), and shaken for four hours. After filtration, the resin was washed with water and, subsequently, extracted three times with 100% methanol (1:1, w:v). The dried crude extract (340.1 mg) was solved in 60% methanol and then subjected to open column chromatography using Polygoprep 60-50 C_{18} (Macherey-Nagel) as stationary phase. The column was successively eluted with 250 mL of 20%, 40%, 60%, 80%, and 100% methanol in water. ^{1}H-NMR spectra were recorded for every fraction. The ^{1}H-NMR spectrum of the 80% fraction showed several signals in the region between 4.00 and 4.50 ppm,

which were assigned to α methine protons of amino acid-derived compounds. Thus, the 80% methanol fraction was further separated by semi-preparative HPLC using a C_{18} column (Nucleodur HTec 250 × 10 mm, 5 µm; Macherey Nagel). The mobile phase consisted of 0.1% trifluoroacetic acid in water (solvent A) and methanol (solvent B). The flow rate was 1.0 mL/min. The gradient conditions were: 0-10 min, 50% to 80% B; 10-30 min, 80%-100% B; 30-40 min, 100% B. The second purification was performed with a C_8 column (Eclipse XDB-C_8, 250 × 9.4 mm, 5 µm; Agilent) using a gradient of acetonitrile (B) in water +0.1% trifluoroacetic acid (A) running from 40% B to 100% B over 50 min with a flow rate of 1 mL min^{-1}. This procedure resulted in the isolation of two compounds, cyclo(Phe-Pro) (**1**) and cyclo(Tyr-Pro) (**2**).

Structure elucidation of cyclo(Phe-Pro) (1)

The molecular formula of cyclo(Phe-Pro) was assigned as $C_{14}H_{16}N_2O_2$ by HR-ESI-MS (*m/z* 245.2788 $[M+H]^+$). The molecular formula was consistent with eight degrees of unsaturation. The ^{13}C NMR spectrum contains eight signals that were assigned to sp^2-hybridized carbons, including two carbonyl groups resonating at δ_C 169.1 (C-5) and δ_C 165.4 (C-7). The other six sp^2-hybridized carbons can be attributed to a benzene ring. Thus, the remaining two unsaturation units required the presence of two ring systems in the compound **1**. Besides, six additional carbon atoms (C-1, C-2, C-3, C-4, C-6 and C-8) are sp^3-hybridized according to their chemical shifts in the high-field region (Table 3). Analysis of the ^1H-NMR and COSY spectra revealed three discrete spin systems, including a monosubstituted benzene, a -CH-CH_2-CH_2-CH_2- moiety and a -NH-CH-CH_2- moiety. The spin systems were assigned to a proline and a phenylalanine residue on the basis of HMBC correlations. Thus, compound **1** was identified as cyclo(Phe-Pro).



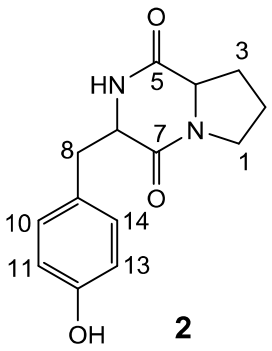
1

Table 3. NMR spectral data for cyclo(Phe-Pro) (**1**) in DMSO-d_6

C-#	δ_C	δ_H, mult., *J* [Hz]		COSY	HMBC
1	44.7	a: 3.26, m b: 3.38, m	CH_2	H-2	C-2, 3
2	22.0	1.73, m	CH_2	H-1, 3	C-1, 3
3	27.9	a: 1.43, m b: 2.01, m	CH_2	H-2, 4	C-1, 2, 4
4	58.6	4.08, m	CH	H-3a, 3b	C-3, 5
5	169.1		C_q		
6	56.3	4.34, m	CH	H-8, -NH*	C-7, 8, 9
7	165.4		C_q		
8	35.6	3.05, dd (5.2, 5.1)	CH_2	H-6	C-6, 7, 9
9	137.1		C_q		
10	130.3	7.23, m	CH	H-11	C-12
11	126.9	7.26, m	CH	H-10, 12	C-9, 13
12	128.2	7.18, m	CH	H-11, 13	C-10, 14
13	126.9	7.26, m	CH	H-12, 14	C-9, 11, 14
14	130.3	7.23, m	CH	H-13	C-12, 13
		7.97, s	-NH	H-6	

Structure elucidation of cyclo(Tyr-Pro) (2)

HR-ESI-MS identified a quasimolecular $[M+H]^+$ ion with an *m/z* of 261.2212, which corresponds to the molecular formula $C_{14}H_{16}N_2O_3$ with eight degrees of unsaturation. The ^{13}C-NMR spectrum showed the presence of eight signals for sp^2-hybridized carbons, including two carbonyl groups: C-5 (δ_C 166.0) and C-7 (δ_C 165.6). COSY data revealed three spin systems in compound **2**. The doublets at δ_H 7.04 (1H, d, *J* = 8.3 Hz) and δ_H 6.64 (1H, d, *J* = 8.3 Hz) arise from an AA'BB' spin system, which represents a *para*-disubstituted benzene ring. In addition, the COSY spectrum revealed the same proline spin system (-CH-CH_2-CH_2-CH_2-) as in compound **1**. A broad signal at δ_H 9.18 in the ^{1}H-NMR spectrum was assigned to a hydroxyl proton. HMBC correlations from H-8 (δ_H 2.91) to C-6 (δ_C 55.7), C-7 (δ_C 166.0), C-9 (δ_C 126.5), C-10 (δ_C 130.5) and C-14 (δ_C 130.5) led to the identification of a



2

tyrosine residue. Compound **2** was therefore assigned as cyclo(Tyr-Pro).

Table 4. NMR spectral data for cyclo(Tyr-Pro) **(2)** in DMSO-d_6

C-#	δ_C	δ_H, mult., *J* [Hz]		COSY	HMBC
1	44.3	a: 3.25, m b: 3.38, m	CH_2	H-2	C-2, 3, 7
2	27.8	a: 1.64, m b: 1.70, m	CH_2	H-1, 3	C-1, 3
3	23.9	a: 1.39, m b: 1.99, m	CH_2	H-2, 4	C-1, 2, 4, 5
4	57.5	4.04, m	CH	H-3	C-3, 5
5	165.6		C_q		
6	55.7	4.23, m	CH	H-8, -NH*	C-7, 8
7	166.0		C_q		
8	34.6	2.91, d (4.8)	CH	H-6	C-6, 7, 9
9	126.5		C_q		
10	130.5	7.04, d (8.3)	CH	H-11	C-12
11	114.4	6.64, d (8.3)	CH	H-10	C-9, 12
12	155.6		C_q		
13	114.4	6.64 d (8.3)	CH	H-14	C-9, C-12
14	130.5	7.04, d (8.3)	CH	H-13	C-12
		7.85, s	NH	H-6	

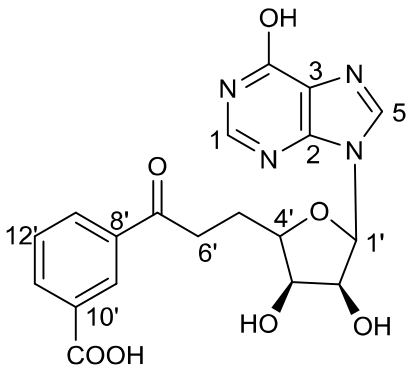
4.2.3 Isolation of secondary metabolites from *H. giganteus*

For chemical analyses, *H. giganteus* DSM 52868 and DSM 52870 were cultured in 2 L shake cultures using VY/2 medium ((baker's yeast, 0.5% (w/v); $CaCl_2$ x 2 H_2O, 0.1% (w/v), pH 7.2). After cultivation at 30 °C and 90 rpm for seven days, the culture broth was centrifuged at 7,000 rpm for 15 min to remove the biomass. The supernatant was extracted three times with ethyl acetate (1:1, v:v). A total of 210.8 and 230.2 mg of crude extract was obtained from DSM 52868 and DSM 52870, respectively. These crude extracts were subjected to open column chromatography using Polygoprep 60-50 C_{18} (Macherey-Nagel) as stationary phase. The two crude extracts were separated into five fractions each

by consecutively increasing the MeOH content of the MeOH-water eluent (20%, 40%, 60%, 80% and 100%, v/v). Since similar signals were present in the ^{1}H-NMR spectra of the 40% fractions from both strains, the two fractions were combined for HPLC purification. The HPLC mobile phase consisted of 0.1% trifluoroacetic acid in water (solvent A) and methanol (solvent B). The gradient program was as follows: 0–10 min, 20–40% B; 10–40 min, 40–100% B; 40–50 min, 100% B; the elution of compounds was monitored at 245, 254, and 280 nm. The flow rate was set to 1.0 mL/min. The compounds futalosine (**3**), 4-*O*-methylfutalosine (**4**) and *N*-acetyltryptophan (**5**) eluted after 34 min, 38 min, and 36 min, respectively.

Structure elucidation of futalosine (3)

Compound **3** possesses the molecular formula $C_{19}H_{18}N_4O_7$ as determined by HR-ESI-MS (*m/z* 415.1251 $[M+H]^+$). The molecular formula indicates 13 degrees of unsaturation. The ^{13}C-NMR spectrum showed the presence of signals for thirteen sp^2-hybridized carbons including two carbonyl groups resonating at δ_C 198.9 and 166.6 (Table 5). In the ^{1}H-NMR spectrum, signals at δ_H 8.43 (H-9', 1H, s), 8.14 (H-11', 1H, d, *J* = 7.9 Hz), 7.63 (H-12', 1H, t, *J* = 7. 9 Hz) and 8.16 ppm (H-13', 1H, d, *J* = 7.9 Hz) were assigned to a 1,3-disubstituted benzene ring based on HMBC and COSY data. In addition, the COSY spectrum revealed a structural fragment corresponding to a ribose moiety. HMBC correlations from H-5 (δ_H 8.28) to C-1', C-2 and C-3 as well as from H-1 (δ_H 8.02) to C-2 and C-3 allowed the assignment of a hypoxanthine ring in this compound. Comparison with literature data suggested that the isolated compound **3** is futalosine [146].



3

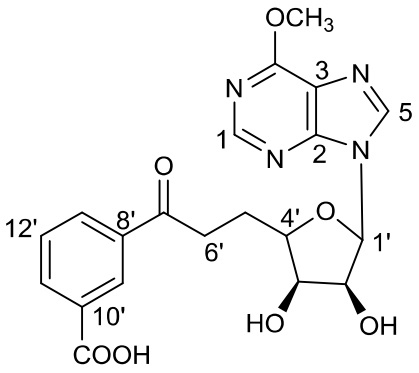
Table 5. NMR spectral data for futalosine (**3**) in DMSO-d_6

C-#	δc	δ_H, mult., *J* [Hz]		COSY	HMBC
1	146.5	8.02, s	CH		C-2, 3
2	147.9		C_q		
3	124.3		C_q		
4	156.5		C_q		
5	136.7	8.28, s	CH		C-1', 2, 3
1'	88.1	5.83, d (5.1)	CH	H-2'	C-3', 2, 5
2'	74.2	4.58, m	CH	H-1', 3'	C-1', 3'
3'	73.2	4.08, t (5.2)	CH	H-2'	C-1', 2', 4'
4'	82.8	3.94, m	CH	H-5'	C-3', 5', 6'
5'	27.0	a: 1.40, m b: 1.99, m	CH_2	H-4'	C-3', 4', 6', 7'
6'	34.3	3.18, m	CH_2	H-5'	C-4', 5', 7'
7'	198.9		C_q		
8'	133.6		C_q		
9'	130.6	8.43, s	CH		C-7', 8', 10', 10'-COOH
10'	131.7		C_q		
11'	136.6	8.14, d (7.9)	CH	H-12'	C-12'
12'	128.6	7.63, t (7.9)	CH	H-11', 13'	C-11', 13'
13'	132.5	8.16, d (7.9)	CH	H-12'	C-12'
10'- COOH	166.6	12.37, s (-OH)			C-9'

Structure elucidation of 4-*O*-methylfutalosine (4)

The molecular formula of **4** was established as $C_{20}H_{20}N_4O_7$ by HR-ESI-MS (*m/z* 429.1279 $[M+H]^+$) indicating the same degree of unsaturation as in **3**. Five oxygenated sp^3 carbon and thirteen sp^2 carbon resonances were observed in the ^{13}C-NMR spectrum (Table 6). The comparison of the ^{1}H and ^{13}C data of compound **4** with those of compound **3** showed that the former also features a 1,3-

disubstituted benzene ring and a ribose moiety. In addition, the HMBC spectrum suggested the presence of a hypoxanthine ring in this compound. The extra carbon atom (δ_C 63.0) in **4** was assigned to a methoxy group. The latter is located at C-4 according to HMBC data. Thus, compound **4** was identified as 4-*O*-methylfutalosine.

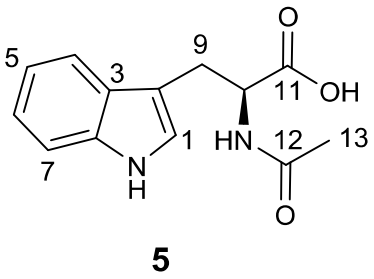


4

Table 6. NMR spectral data for 4-*O*-methylfutalosine (**4**) in DMSO-d_6

C-#	δ_C	δ_H, mult., *J* [Hz]		COSY	HMBC
1	151.8	8.03, s	CH		C-2, 4
2	150.6		C_q		
3	130.8		C_q		
4	156.8		C_q		
5	136.7	8.20, s	CH		C-2, 3
1'	86.4	5.92, d (7.6)	CH	H-2'	C-2', 2, 5
2'	70.4	4.94, dd (5.49, 5.56)	CH	H-1', 3'	C-1', 3'
3'	76.0	4.23, t (5.5)	CH	H-2'	C-2'
4'	83.7	3.91, m	CH	H-5'	C-5'
5'	30.1	2.15, m	CH_2	H-4', 6a', 6b'	C-4', 6'
6'	25.0	a: 2.74, m b: 3.02, m	CH_2	H-5'	C-5', 7'
7'	196.5		C_q		
8'	134.1		C_q		
9'	127.0	8.38, s	CH		C-8', 10'-COOH
10'	134.2		C_q		
11'	136.2	8.19, d (8.0)	CH	H-12'	C-10', 12', 13'
12'	127.0	7.72, m	CH	H-11', 13'	C-10', 11', 13'
13'	131.0	7.96, d (7.7)	CH	H-12'	C-8', 11', 12'
10'-COOH	166.6		C_q		C-9'
-OCH_3	63.0	3.33, s	CH_3		C-4

Structure elucidation of *N*-acetyltryptophan (5)



5

The chemical formula of compound **5** was established as $C_{13}H_{14}N_2O_3$ by HR-ESI-MS (*m/z* 247.1026 $[M+H]^+$), which indicates eight degrees of unsaturation. The ^{13}C-NMR spectrum showed the presence of ten signals for sp^2-hybridized carbon atoms including two carbonyl groups (Table 7). Analysis of ^{1}H-NMR and COSY spectra revealed an aromatic spin system consistent with an *ortho*-disubstituted benzene moiety. The presence of an indole ring was supported by a COSY cross-peak between an exchangeable proton at δ_H 10.82 and H-1. The same exchangeable proton gave heteronuclear long-range correlations to C-1, C-2, C-3, and C-8. Further interpretation of HMBC data revealed a tryptophan residue, which is acetylated at the α-amino group.

Table 7. NMR spectral data for *N*-acetyl-tryptophan (**5**) in DMSO-d_6

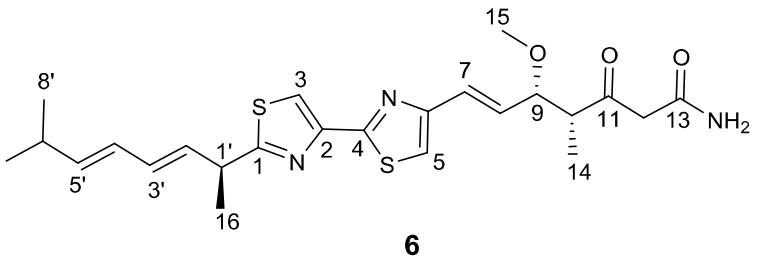
C-#	δ_C	δ_H, mult., *J* [Hz]		COSY	HMBC
1	123.7	7.13, s	CH		C-2, 3
2	109.7		C_q		
3	127.2		C_q		
4	118.0	7.52, d (8.2)	CH	H-5	C-6
5	120.4	6.97, m	CH	H-4, 6	C-3, 7
6	121.0	7.05, m	CH	H-5, 7	C-4, 8
7	111.6	7.33, d (8.2)	CH	H-6	C-5
8	136.1		C_q		
9	27.1	2.93, m	CH_2	H-10	C-2, 11
10	52.9	4.47, m	CH	H-9	C-9, 11
11	173.6		C_q		
12	169.2		C_q		
13	22.4	1.79, s	CH_3		C-12
		10.82, s	-NH	H-1	C-1, 2, 3, 8
		8.13, s	-NH	H-10	C-9, 10, 12

4.2.4 Isolation of desmethyl-myxothiazol from *Archangium minus* DSM 52715

The myxobacterium *Archangium minus* DSM 52715 was obtained from the DSMZ and cultured in 5 L VY/2 medium for five days at 150 rpm and 30 °C. The extraction and fractionation followed the protocol described in chapter 4.2.5. The different fractions obtained from the open column chromatography were subjected to a disk diffusion assay [147]. The samples were dissolved in methanol (1 mg/mL) and then 50 µL of each sample were transferred to a single paper disc. Ciprofloxacin and amphotericin B were used as positive controls. In this assay, the 100% methanol fraction exhibited the broadest activity spectrum against Gram-positive and Gram-negative bacteria as well as fungi. Thus, the corresponding extract was further purified by semipreparative HPLC using a C_{18} column (Nucleodur HTec 250 × 10 mm, 5 µm; Macherey Nagel) with a linear gradient of acetonitrile (B) in water plus 0.1% trifluoroacetic acid (A): 0–30 min, 70–100%, B; 30–50min, 100% B. The elution of metabolites was monitored at 245, 254, and 280 nm. The flow rate was set to 1.0 mL/min. This procedure resulted in the isolation of compound **6**.

Structure elucidation of desmethyl-myxothiazol (6)

Compound **6** showed a pseudomolecular ion at *m/z* 475.1883 $[M+H]^+$ in the HR-ESI-MS, which was consistent with a chemical formula of $C_{24}H_{31}N_3O_3S_2$ (calcd for 475.6613) and eleven degrees of unsaturation. The ^{13}C-NMR spectrum revealed fourteen signals, which could be ascribed to sp^2-hybridized carbons based on their chemical shifts (Table 8). Analysis of COSY and ^{1}H-NMR spectra revealed two independent spin systems, a $CH(CH_3)_2$-CH-CH-CH-CH-$CH(CH_3)$- and a -CH-CH-CH(OR)-$CH(CH_3)$ moiety. From the HMBC spectrum, correlations from H-3 to C-1, 2 and 4, from H-5 to C-4, 6 and 7, from H-2' to C-1 and from H-8 to C-6 established the linkage of these two spin systems. The HMBC data also indicated the location of a methoxy group at C-9. A sp^2-hybridized



6

carbon atom at δ_C 176.9 (C-1) was assigned to be part of a 2,4-disubstituted five-membered ring as evidenced on the correlation of the aromatic methine H-3 (δ_H 8.04, s) to C-1 and another quaternary sp^2 carbon C-2 (δ_C 148.3) in the HMBC spectrum.

Table 8. NMR spectral data for desmethyl-myxothiazol (**6**) in DMSO-d_6

C-#	δ_C	δ_H, mult., *J* [Hz]		COSY	HMBC
1	176.9		C_q		
2	148.3		C_q		
3	115.9	8.04, s	CH		C-1, 2
4	163.2		C_q		
5	116.5	7.43, s	CH		C-4, 6
6	153.7		C_q		
7	126.4	6.71, d (15.3)	CH	H-8	C-6, 9
8	129.4	6.45, dd (7.6, 15.3)	CH	H-7, 9	C-6, 9
9	82.6	4.11, dd (5.3, 7.3)	CH	H-8, 10	C-7, 8, ,11, 14, 15
10	50.6	3.09, m	CH	H-9, 14	C-8, 9, 14
11	206.4		C_q		
12	40.4	3.5, s	CH_2		C-10, 11, 13
13	177.9		C_q		
14	10.6	1.14, d (7.1)	CH_3	H-10	C-9, 11
15	55.4	3.35, s	CH_3		C-9
16	19.6	1.55, d (7.2)	CH_3	H-1'	C-1
1'	41	3.94, m	CH	H-2', 16	C-1, 2', 3'
2'	130.1	5.79, dd (7.9, 15.7)	CH	H-1', 3'	C-1', 3', 4', 16
3'	131.0	6.23, dd (10.5, 15.3)	CH	H-2', 4'	C-1', 2', 4', 5'
4'	126.7	6.05, dd (10.5, 15.3)	CH	H-3', 5'	C-3'
5'	141.8	5.67, dd (5.7, 14.3)	CH	H-4', 17	C-3', 6'
6'	30.8	2.35, m	CH	H-7', 8'	C-5'
7', 8'	21.2	1.04, m	CH_3		

Additional results

Another five-member ring was established according to the correlation of H-5 (δ_H 7.43, s) to two quaternary sp^2 carbons. The long-range correlations in HMBC as well as the observed chemical shifts were similar to those of the thiazole rings in myxothiazols. A comparison with published data suggested that **6** is desmethyl-myxothiazol. This compound and its derivatives are commonly reported from myxobacteria [147-148].

5 Discussion

5.1 Genome mining for biosynthetic gene clusters in the phylum *Chloroflexi*

Among 43 specified species in the phylum *Chloroflexi*, genome sequences of 35 strains were published and released to date. Mining of these complete genome sequences using bioinformatics tools, such as antiSMASH, Pfam and Bagel3, allowed the identification of gene clusters associated with the production of secondary metabolites. Similar to other bacterial phyla [97], genes encoding the biosynthesis of isoprenoids are widely distributed in the phylum *Chloroflexi*. Among all *Chloroflexi* bacteria, NRPS, PKS and hybrid PKS-NRPS gene clusters are mainly present in the genera *Herpetosiphon* and the species *Ktedonobacter racemifer*. Products of NRPS and PKS gene clusters represent an important class of natural products with a wide range of biomedical and biocontrol potential. Since several secondary metabolites biosynthetic gene clusters from the three species showed no homologous with known clusters, they have potential to produce novel natural products. Thus, *Herpetosiphon* spp. and *K. racemifer* could be proposed as new source of natural product. Bacteria containing genome size smaller than 3 or 4 Mb tend not to produce secondary metabolites. However, the biosynthetic potential of other members of the phylum *Chloroflexi* are difficult to determine due to the limited genomic data.

5.1.1 Isoprenoid biosynthesis in the phylum *Chloroflexi*

Genome sequence analyses have revealed the distribution of the MEP and MEV pathways in the phylum *Chloroflexi*. The five classes *Chloroflexia*, *Anaerolineae*, *Caldilineae*, *Thermoflexia*, and *Ardenticatenia* contain the genes of the MEV pathway, while the three classes *Thermomicrobia*, *Ktedonobacteria*, and *Dehalococcoidia* were found to utilize the MEP pathway for the biosynthesis of isoprenoids. To reveal production of isoprenoid products in the phylum *Chloroflexi*, genome-wide analyses were conducted using hidden Markov models (HMMs) and protein families database (Pfam)

searches. Analysis of 48 genome sequences from classified and unclassified *Chloroflexi* bacteria led to the discovery of 42 terpene synthases. Among these synthases, 25 belong to the family of phytoene synthase that are widely distributed in the phylum *Chloroflexi*. Sesquiterpene and diterpene synthases occurs mainly in the genus *Herpetosiphon*. In addition, terpene synthases from bacteria in the genus *Herpetosiphon* exhibit highly sequence identities, suggesting that they may contribute to product the same or similar isoprenoid products.

The potential of the *Chloroflexi* for the production of isoprenoids is further supported by chemical analyses. A diterpene named verrucosan-2β-ol was already reported in 1993 from *Chloroflexus aurantiacus* [149]. A heterologous expression experiment resulted in the discovery of *O*-methylkolavelool from *Herpetosiphon aurantiacus* DSM 785 [95]. In the present study, it was shown that isotope labeling is another promising approach to trace as yet undetected isoprenoids. In this way, the new diterpene herpetopanone was discovered from *H. aurantiacus* DSM 785 [150].

Isoprenoids constitute the largest class of natural products and they exert a number of important biological functions. Ubiquinones and menaquinones which feature isoprenoid side-chains are involved in electron transport. Bactoprenols act as cell membrane carrier molecules that are involved in the biosynthesis of peptidoglycan. Carotenoids perform two important functions in photosynthesis: photo-protection and light harvesting [151]. Several non-phototrophic and phototrophic bacteria in the phylum *Chloroflexi* are able to produce carotenoids according to the wide distribution of phytoene synthase genes in this phylum. Phytoene synthases catalyze the formation of phytoene, which is an intermediate in the biosynthesis of carotenoids. Aside from their role in photosynthesis, carotenoids may also act as antioxidant and photoprotection agents. Biological activities of the three isoprenoids verrucosan-2β-ol, *O*-methylkolavelool and herpetopanone have not been described yet. The structure of herpetopanone features an octahydro-1*H*-indenyl skeleton, which is also known from the sesquiterpene oplopanone [152]. The latter was isolated from several medicinal plants, such as *Alpinia oxyphylla*, *Reneilmia cincinnata* and *Homalomena sagittifolia* [153-155]. Oplopanone possesses week activity against the human malaria parasite *Plasmodium falciparum* [154]. Due to its structural similarity, herpetopanone may have anti-plasmodial activity as well.

5.1.2 Natural products biosynthesis in the genus *Herpetosiphon*

The *Herpetosiphon* genomes contain a high number of NRPS, PKS and hybrid PKS/NRPS biosynthetic loci similar to other wolf pack-forming bacteria, such as myxobacteria and *Lysobacter* species. Predation in these predatory bacteria were proposed to depend on the production of bioactive compounds. However, the amount of secondary metabolites discovered from the genus *Herpetosiphon* was less than our expected based on genomic analyses. *H. aurantiacus* DSM 785 and *H. geysericola* 7119 contain common clusters encoding the biosynthesis of isoprenoids, siderophore, alkyl pyrones, and other unknown products. A putative gene cluster from the two species involved in the biosynthesis of myxochelin-type siderophores. The myxochelins were originally discovered in myxobacteria, showing antitumor activity [114, 156]. Type III PKSs from *H. aurantiacus* and *H. geysericola* show high phylogenetic identities with an alkylpyrone synthase from *Bacillus subtilis*. These common metabolites play roles as agent of photosynthesis protection (carotenoid), signaling, growth regulatory (isoprenoids), and iron chelating (siderophore), as well as membrane composition (alkylpyrones). Only four unique gene clusters were detected in *H. aurantiacus,* including a NRPS, a type I PKS and two hybrid PKS/NRPS gene clusters. These biosynthetic gene clusters are involved in the biosynthesis of strain-specific metabolites, providing information about the chemical diversity and possible events of horizontal gene transfer during evolution in the genus *Herpetosiphon*.

5.1.3 Natural product biosynthesis in the genera *Ktedonobacter* and *Thermogemmatispora*

Together with species in the genus *Thermogemmatispora*, *Ktedonobacter racemifer* belonging to the genus *Ktedonobacter* was placed in the class *Ktedonobacteria*. The size of genome of *Ktedonobacter racemifer* DSM 44963 is quite large (13.7 Mb), which encodes genes for the biosynthesis of resistomycin analogs and other compounds with diverse structures. The high homology of the putative resistomycin biosynthetic gene cluster from *K. racemifer* to that of *Streptomyces resistomycificus* [157] suggested a common ancestor of the two clusters. Thus, the detection of resistomycin biosynthesis genes in distant taxonomic lineages, such as *Streptomyces* and

Ktedonobacter, providing evidence of horizontal gene transfer as orthologs can be found in these lineages. While the production of resistomycin analogs has not been reported from *K. racemifer* to date. The metabolically of bacteria in the class *Ktedonobcteria* is highly diverse. NRPS and PKS gene clusters show a nonuniform distribution in the class *Ktedonobacteria*. Unlike *Ktedonobacter*, no PKS genes were detected in the two species of the genus *Thermogemmatispora*. The NRPS gene clusters from *T. onikobensis* NBRC 111776 and *T. carboxidivorans* DSM 45816 encode the biosynthesis of different compounds. This means that they are not conserved.

5.2 Identification and taxonomy of new *Herpetosiphon* species

When this PhD project was started, the genus *Herpetosiphon* included only two species and it was unclear whether this limited diversity actually represented the situation in nature. There are several examples of taxonomic classifications that are based on metabolic differences [158-159]. For this reason, the identification of novel species might also reveal untapped chemical resources. Thus, the description of two species in the genus *Herpetosiphon*, *H. gulosus* and *H. giganteus*, did not only increase species richness, but also provided opportunities to investigate possible correlations of phylogenetic diversity and secondary metabolite biosynthesis in this genus. Based on genomics analysis of *H. aurantiacus* DSM 785 and *H. geysericola* DSM 7119 (manuscript in preparation), *H. gulosus* and *H. giganteus* could be also interesting secondary metabolite producers. The genomes of these two species will be sequenced in the future.

5.3 Predatory behavior in *Herpetosiphon* spp.

Predatory bacteria exhibit characteristic prey ranges. For example, the potential therapeutic agents, *Bdellovibrio* spp. and *Micavibrio* spp. feed on Gram-negative bacteria; myxobacteria prey on a wide range of Gram-negative and Gram-positive bacteria, including several human pathogens. It was proposed that secondary metabolite production and hydrolytic enzyme secretion depend on nutrient conditions [75]. All *Herpetosiphon* species were reported to be capable of lysing Gram-negative and Gram-positive bacteria, suggesting that predatory behavior is a common trait in this genus. Among

all *Herpetosiphon* species, *H. gulosus* was found to exhibit the strongest predation activity. Interestingly, *H. geysericola* DSM 7119 shows faster swarm expansion on all prey organisms than *H. aurantiacus* DSM 785, although it has a smaller genome sizes and less biosynthetic gene clusters than this strain. The different predation ability among *Herpetosiphon* species may be caused by the genetic diversity, which led to the utilization of distinct bioactive secondary metabolites or lytic enzymes. A better understanding of the predatory behavior in *Herpetosiphon* species may be applied to improve predatory ability and the production of secondary metabolites. In addition, predation assay could be served as a mean to activate cryptic biosynthetic cluster.

5.4 Natural product discovery from wolf-pack strategy predators

Among all predatory bacteria, predation strategy in myxobacteria, lysobacteria and *Herpetosiphon* spp. are referred to 'wolf-pack' or 'group' attack. Chemical analyses of *Myxococcus xanthus* indicated the correlation of antibiotic production and predatory behavior. At present, several natural products with interesting biological activities and chemical structure have been discovered from 'wolf-pack' strategy predators: myxobacteria and lysobacteria. Thus, predatory bacteria using this strategy was considered as promising sources of antibiotics. Their large biosynthetic potential further confirmed this assumption. Compared to myxobacteria and lysobacteria, *Herpetosiphon* species contain less secondary metabolite gene clusters. While genomic analysis revealed that novel natural products can be produced by these strains. Thus, searching secondary metabolites in *Herpetosiphon* spp. would provide evidences to investigate the distribution of metabolic traits among predatory bacteria.

5.4.1 Natural product discovery from *Herpetosiphon*

Genome mining of *H. aurantiacus* DSM 785 and *H. geysericola* DSM 7119 indicated that more than 80% of their secondary metabolites await identification. In addition, approximately 70% of gene clusters from *H. aurantiacus* and *H. geysericola* could not be associated with known pathways. The discovery of new strains showing different predation abilities have potential to biosynthesize unique

secondary metabolites. Thus, there is a great potential for the discovery of new natural products from the genus *Herpetosiphon*. Although *H. geysericola* DSM 7119 and *H. aurantiacus* DSM 785 share a diverse range of secondary metabolite biosynthetic gene clusters, which encode the biosynthesis of isoprenoids (*O*-methylkolavelool, herpetopanone and carotenoid), nonribosomal peptide (myxochelin A) and polyketides (alkylprones). No secondary metabolites have been isolated from *H. geysericola* due to its slow-growing under the common laboratory condition.

By use of chemical analysis in *H. aurantiacus* and *H. giganteus*, five known compounds were isolated from these bacteria. Two diketopiperazines, cyclo (Phe-Pro) and cyclo (Tyr-Pro) were discovered from *H. aurantiacus*. These two diketopiperazines have been commonly found in several bacteria, showing antifungal and antibacterial activities. Since biological activities of herpetopanone, *O*-methylkolavelool and auriculamide are unknown, these two diketopiperazine as bioactive compounds detected for the first time in *H. aurantiacus*. Their roles in predation remain unclear. Chemical analysis of culture broth of *H. giganteus* led to the isolation of futalosine and its derivatives, as well as *N*-acetyl-tryptophan from this bacterium. Futalosine was previously reported from *Streptomyces* sp., serving as biosynthetic intermediate in ubiquinone pathway. At present, the ubiquinone pathway was only found in alpha-, beta-, gamma-proteobacteria. Thus, the presence of this compound proposed that the ubiquinone pathway might be utilized in *Herpetosiphon* sp. for electron transport. In addition, LC-MS analysis of crude extract of *H. aurantiacus* showed a mass shift which would dedicated to a diterpene based on the molecular formula. The usage of an isotope labeling approach resulted in the isolation of this diterpene herpetopanone from *H. aurantiacus* DSM 785. Although isotope labeling approach was widely used in actinobacteria and myxobacteria for natural product isolation, isoprenoid with octahydroindene skeleton discovered from native producer for the first time by this approach. While secondary metabolites reported from *H. aurantiacus* or *H. giganteus* have not been detected in other *Herpetosiphon* spp.. Differences in chemical profiles between *H. aurantiacus* and *H. giganteus* are caused by many reasons, such as different culture conditions, biodiversity, and inefficient process of natural product isolation.

At present, the exploration of the chemistry of *Herpetosiphon* spp. faces many difficulties and

problems, such as low production levels, gene non-expression and inefficient isolation processes. Traditionally, bioassay-guided fractionation was performed in the search for new secondary metabolites. However, this approach bears the risk of rediscovering known compounds. Furthermore, it is difficult to select the correct bioassay for the identification of structurally novel metabolites. Considering that most of undetected secondary metabolites in *Herpetosiphon* spp. are polyketides, non-ribosomal peptides and ribosomal peptides. OSMAC and heterologous expression strategies are worthwhile applying for production improvement of secondary metabolites and silent genes activation in future research. While three hybrid PKS-NRPS clusters spanning over 50 kb from *H. aurantiacus* are extremely difficult to express in heterologous expression systems. On the use of OSMAC approach by varying the source of carbon in the growth medium was fail in this project to activate cryptic biosynthetic pathways in *H. aurantiacus*. Searching a useful recombination system for heterologous expression and adequate culture conditions would be quite important in future natural product research in the genus *Herpetosiphon*. In addition, the strongest predation ability of *H. gulosus* excited our interest for the metabolite production based on the assumption that bioactive compounds play critical role in predation. The biosynthetic potential of *H. gulosus* have not been investigated in this project due to the time restriction. It would be evaluated in the future.

5.4.2 Natural product discovery from *Archangium minus*

At present, myxobacterial taxa contains 62 validly described species to date. Myxobacterium *Archangium minus* was transferred from the genus *Cystobacter* into a genus *Archangium* in 2015. Myxobacteria is best known as bioactive compound producer, whereas no secondary metabolites are reported from this bacterium to date. Chemical analysis of the culture of *Archangium minus* allowed the isolation of desmethyl-myxothiazol in this project. Desmethyl-myxothiazol was detected from several myxobacteria, showing antibacterial activity. The presence of this compound indicated that this species is capable of producing bioactive product. Desmethyl-myxothiazol could be used as reference to reveal the correlation chemical diversity with taxonomic distance.

Up to now several secondary metabolite biosynthetic gene clusters are waiting to be revealed their

encoded products. The increasing of bacterial genome sequence will cast new light on the distribution of biosynthetic pathways among the phylum *Chloroflexi*. Besides, along with technological development will provide the possibility to awake cryptic biosynthetic gene clusters and thus to exploit bacterial metabolomes. The metabolic diversity of *Herpetosiphon* spp. will be unveiled in the future rely on these advances.

Summary

Bacteria with predatory behavior are found in the well-known phyla *Actinobacteria* and *Proteobacteria*, which include many proliferative producers of antibiotics. Furthermore, previous studies have confirmed that myxobacteria utilize bioactive secondary metabolites as part of their predation strategy. Genomic analyses have revealed a high number of biosynthetic gene clusters in some predatory bacteria. It is assumed that the biosynthesis of antibiotics plays an important role in their predatory lifestyle.

Herpetosiphon was classified into the phylum *Chloroflexi* and is hence only distantly related to other predatory bacteria. In comparison to other bacteria of the phylum *Chloroflexi*, *Herpetosiphon* spp. were found to possess the highest genomic potential for natural product biosynthesis. A total of 14 and 9 biosynthetic gene clusters were identified from *H. aurantiacus* DSM 785 and *H. geysericola* DSM 7119, respectively. Even after the discovery of auriculamide and *O*-methylkolavelool, more than 80% of the secondary metabolome of *H. aurantiacus* DSM 785 is uncharted.

Chemical analyses of extracts from different *Herpetosiphon* species led to the isolation of five known metabolites, including the two diketopiperazines cyclo(Phe-Pro) and cyclo(Tyr-Pro), the nucleosides futalosine, and *O*-methylfutalosine, as well as *N*-acetyl-tryptophan. The role of these compounds in predation is still unclear. A new natural product was discovered from *H. aurantiacus* DSM 785 after a feeding experiment with ^{13}C-labeled glucose. Following the diversion of the ^{13}C label into the MEP pathway, the diterpene herpetopanone was traced due to its characteristic isotope pattern in MS analyses. The compound obscuronatin, which had previously been reported after the heterologous expression of a diterpene synthase from *H. aurantiacus* DSM 785 might actually represent an intermediate in the biosynthesis of herpetopanone. An investigation on the biological activity of this compound is in progress.

In this thesis, a new species named *H. gulosus* was identified using genetic and phenotypic analyses. *H. gulosus* showed an exceptionally high predation activity against Gram-negative and Gram-positive bacteria. In addition, lawn predation assay revealed that all *Herpetosiphon* species exhibited varying

degrees of predation activity against prey organisms. The observed differences in the predation activities may be related to the production of different antibiotics and lytic enzymes.

A comparison of genome sequences from 35 *Chloroflexi* bacteria revealed that species from the genera *Herpetosiphon*, *Thermogemmatispora* and *Ktedonobacter* possess a huge potential for secondary metabolite production. Similar to *Herpetosiphon* species, a variety of secondary metabolite biosynthetic gene clusters were identified from the genome of *Ktedonobacter racemifer* DSM 44963. The corresponding loci were predicted to govern the biosynthesis of isoprenoids, polyketides, nonribosomal and ribosomal peptides. Of particular interest was the finding of a type II PKS gene cluster, which might be involved in the production of resistomycin-type compounds. This type of biosynthesis was previously only known from *streptomycetes*. Although the two strains *T. onikobensis* NBRC 111776 and *T. carboxidivorans* DSM 45816 are unable to produce polyketides, unprecedented NRPS and RiPP biosynthetic gene clusters were identified in their genomes. The structural prediction of the encoded compounds might help to design appropriate fermentation and isolation procedures and thus promote future research on the chemistry of these organisms.

Zusammenfassung

Räuberische Bakterien sind sowohl in dem Phylum der *Actinobacteria*, als auch in dem der *Proteobacteria* zu finden. Sie erweisen sich als vielversprechende Produzenten von Antibiotika. Aktuelle Studien belegen, dass die Produktion von bioaktiven Naturstoffen in Myxobakterien ein Teil ihres räuberischen Verhaltens darstellt. Durch die Genom-Analyse konnte zudem eine auffallend hohe Dichte an Biosynthese Genclustern in räuberischen Bakterien nachgewiesen werden. Dies bestätigt die Annahme, dass die Biosynthese von Antibiotika eine besondere Rolle im räuberischen Lebensstil dieser Bakterien einnimmt.

Herpetosiphon, ein Vertreter des Phylums *Chloroflexi*, besitzt einen weit entfernten Verwandtschaftsgrad zu anderen räuberischen Bakterien. Verglichen mit den genomischen Daten anderer Bakterien des Phylums *Chloroflexi* besitzt *Herpetosiphon* spp. das größte Potential für die Bildung von Naturstoffen. In *H. aurantiacus* DSM 785 wurden insgesamt 14, und in *H. geysericola* DSM 7119 9 Biosynthese Gencluster identifiziert. Trotz der Aufreinigung von Auriculamid und *O*-Methylkolavelool konnten über 80 % der Sekundärmetabolite von *H. aurantiacus* DSM 785 bisher noch nicht isoliert werden.

Chemische Analysen der Extrakte von verschiedenen *Herpetosiphon* Spezies führten zu der Isolierung von fünf Metaboliten, welche bereits in der Literatur beschrieben wurden. Dazu zählen die Cyclo-(Phe-Pro)- sowie die Cyclo-(Tyr-Pro)-diketopiperazine, die Nukleoside Futalosin und *O*-Methylfutalosin, sowie *N*-Acetyl-tryptophan. Ob diese Metabolite ebenfalls eine Rolle in dem räuberischen Verhalten von *Herpetosiphon* spielen ist noch unklar.

Nach der Fütterung von *H. aurantiacus* DSM 785 mit ^{13}C-markierter Glukose konnte ein neuer Naturstoff gefunden werden. Gemäß der Weiterleitung der ^{13}C-markierten Glukose in den MEP-Biosyntheseweg, konnte anschließend das Diterpen Herpetopanon an Hand seines charakteristischen Isotopen-Musters in der MS-Analyse identifiziert werden. Die Verbindung Obscuronatin, welche vor kurzem in der Literatur beschrieben wurde nachdem sie mittels heterologer Expression einer Diterpen-Synthase von *H. aurantiacus* DSM 785 hergestellt wurde,

stellt ein putatives Intermediat in der Biosynthese von Herpetopanon dar. Die genaue biologische Rolle von Obscuronatin wird aktuell noch untersucht.

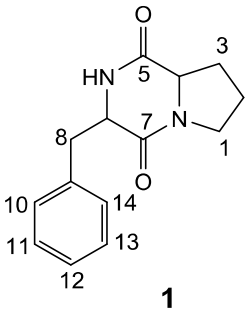
In der vorliegenden Arbeit wurde eine neue Spezies, *H. gulosus*, mittels genetischer und phänotypischer Analysen identifiziert. Das räuberische Verhalten von *H. gulosus* war besonders gegenüber grampositiven und gramnegativen Bakterien ausgeprägt. Unter Verwendung des sogenannten „lawn predation assay“, bei dem die Beuteorganismen ähnlich einem Rasen gleichmäßig auf dem Agar wachsen, wurde ersichtlich, dass sich die *Herpetosiphon* Arten untereinander hinsichtlich ihres Jagdverhaltens unterscheiden lassen. Diese merklichen Unterschiede in dem Ausmaß an räuberischer Aktivität scheinen in Relation zu der Anzahl an produzierten Antibiotika und lytischen Enzymen zu stehen.

Die Genomsequenzen von 35 Bakterien des Phylums *Chloroflexi* wurden miteinander verglichen und dabei wurde ersichtlich, dass besonders Vertreter der Gattungen *Herpetosiphon*, *Thermogemmatispora* und *Ktedonobacter* über ein enormes Potential zur Biosynthese von Naturstoffen verfügen. Das Genom von *Ktedonobacter racemifer* DSM 44963 besitzt, ähnlich zu den Genomen der *Herpetosiphon* Arten, eine Vielzahl von Genclustern, welche für die Biosynthese von unterschiedlichsten Sekundärmetaboliten kodieren. Die identifizierten Genloki sind für die Biosynthese von Isoprenoiden, Polyketiden, nichtribosomalen und ribosomalen Peptiden verantwortlich. Von besonderer Interesse war hierbei ein Typ II PKS Genlcuster, welches möglicherweise an der Biosynthese von Resistomycin-artigen Verbindungen beteiligt ist. Diese Biosynthese, war bisher nur aus Streptomyceten beschrieben worden. *T. onikobensis* NBRC 111776 und *T. carboxidivorans* DSM 45816 können keine Polyketide produzieren, jedoch wurden völlig neuartige NRPS und RiPP Biosynthese-Gencluster in ihren Genomen gefunden. Eine Vorhersage über die möglichen chemischen Strukturen der kodierten Verbindungen ist hilfreich für die Planung der geeigneten Fermentations- und Isolationsbedingen. Zugleich würden diese Informationen zur Verbesserung der zukünftigen Forschung an der Chemie dieser Organismen beitragen.

Appendix

NMR data

All spectra were referenced internally using the deuterated solvent (dimethyl sulfox-d_6: δ_H 2.50 ppm, δ_C 39.5).



1

cyclo(Phe-Pro) (1): colorless oil: ^{1}H-NMR (600 MHz, DMSO-d_6) δ_H [ppm] (J[Hz]) 1.43 (2H, m, Ha-3), 1.73(2H, m, H-2), 2.01(2H, m, Hb-3), 3.05(2H, dd, J 5.2, 5.1, H-6), 3.26 (2H, m, Ha-1), 3.38(2H, m, Hb-1), 4.08(1H, m, H-4), 4.34(1H, m, H-6), 7.18(1H, m, H-12), 7.23(1H, m, H-10), 7.23(1H, m, H-14), 7.26(1H, m, H-11), 7.26(1H, m, H-13), 7.97(1H, s, -NH). ^{13}C-NMR (600 MHz, DMSO-d_6) δ_C [ppm] 22.0 (C-2), 27.9 (C-3), 35.6 (C-8), 44.7 (C-1), 56.3 (C-6), 58.6 (C-4), 126.9 (C-11), 126.9 (C-13), 128.2 (C-12), 130.3 (C-10), 130.3 (C-14), 137.1 (C-9), 165.4 (C-7), 169.1 (C-5).

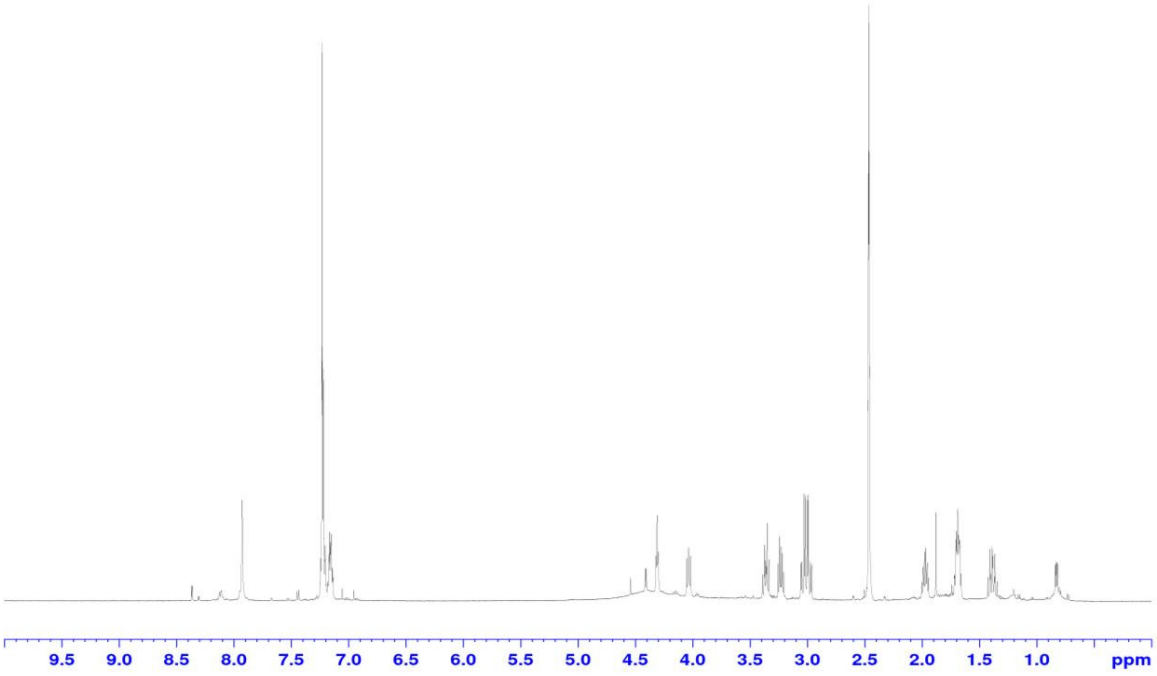


Figure S1. ^{1}H-NMR spectrum of (**1**) in DMSO-d_6.

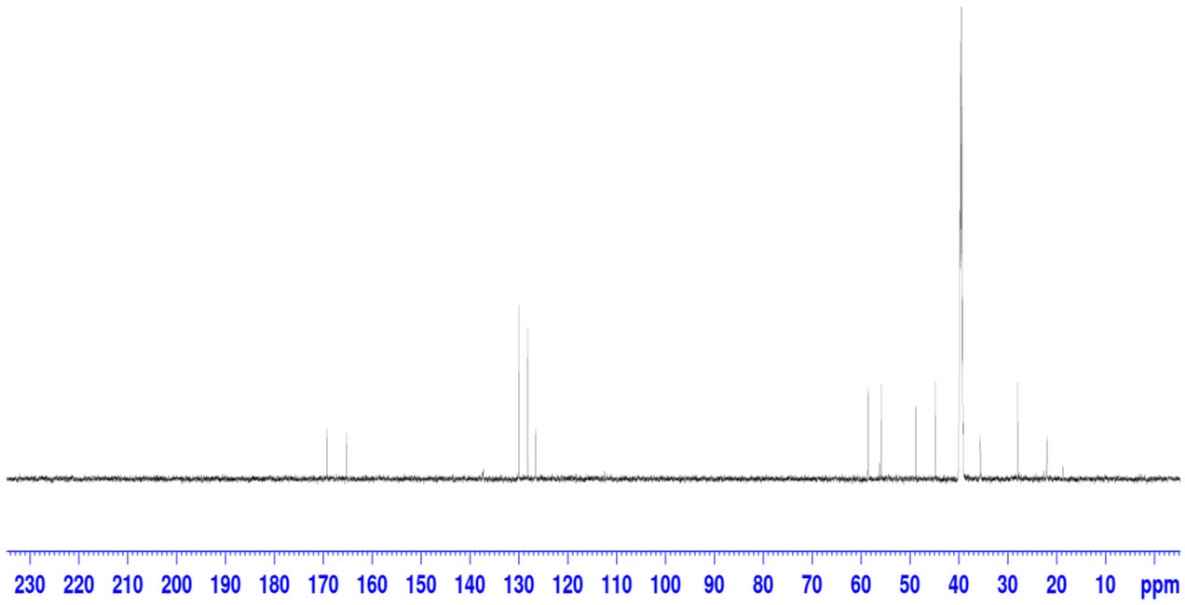
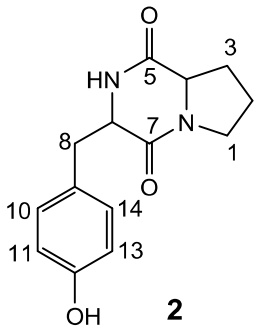


Figure S2. ^{13}C-NMR spectrum of (**1**) in DMSO-d_6.

cyclo(Tyr-Pro) (2): colorless oil: ^{1}H-NMR (600 MHz, DMSO-d_6) δ_H [ppm] (*J*[Hz]) 1.39 (2H, m, Ha-3), 1.64 (2H, m, Ha-2), 1.70 (2H, m, Hb-2), 1.99 (2H, m, Hb-3), 2.91 (H, d, *J* 4.8, H-8), 3.25 (2H, m, Ha-1), 3.38(2H, m, Hb-1), 4.04 (1H, m, H-4), 4.23 (1H, m, H-6), 6.64 (1H, d, *J* 8.3, H-11), 6.64 (1H, d, *J* 8.3, H-13), 7.04 (1H, d, *J* 8.3, H-10), 7.04 (1H, d, *J* 8.3, H-14), 7.85 (1H, s, -NH). ^{13}C-NMR (600 MHz, DMSO-d_6) δ_C [ppm] 23.9 (C-3), 27.8 (C-2), 34.6 (C-8), 44.3 (C-1), 55.7 (C-6), 57.5 (C-4), 114.4 (C-11), 114.4 (C-13), 126.5 (C-9), 130.5 (C-10), 130.5 (C-14), 155.6 (C-12), 165.6 (C-5), 166.0 (C-7).



2

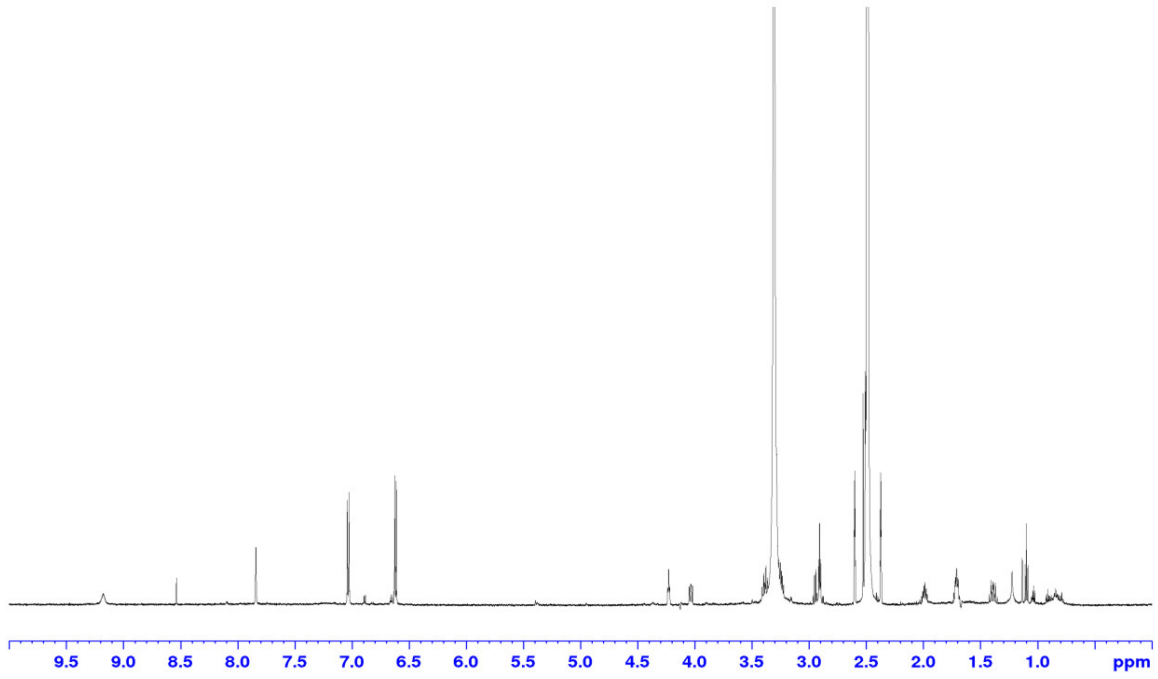


Figure S3. ^{1}H-NMR spectrum of (**2**) in DMSO-d_6.

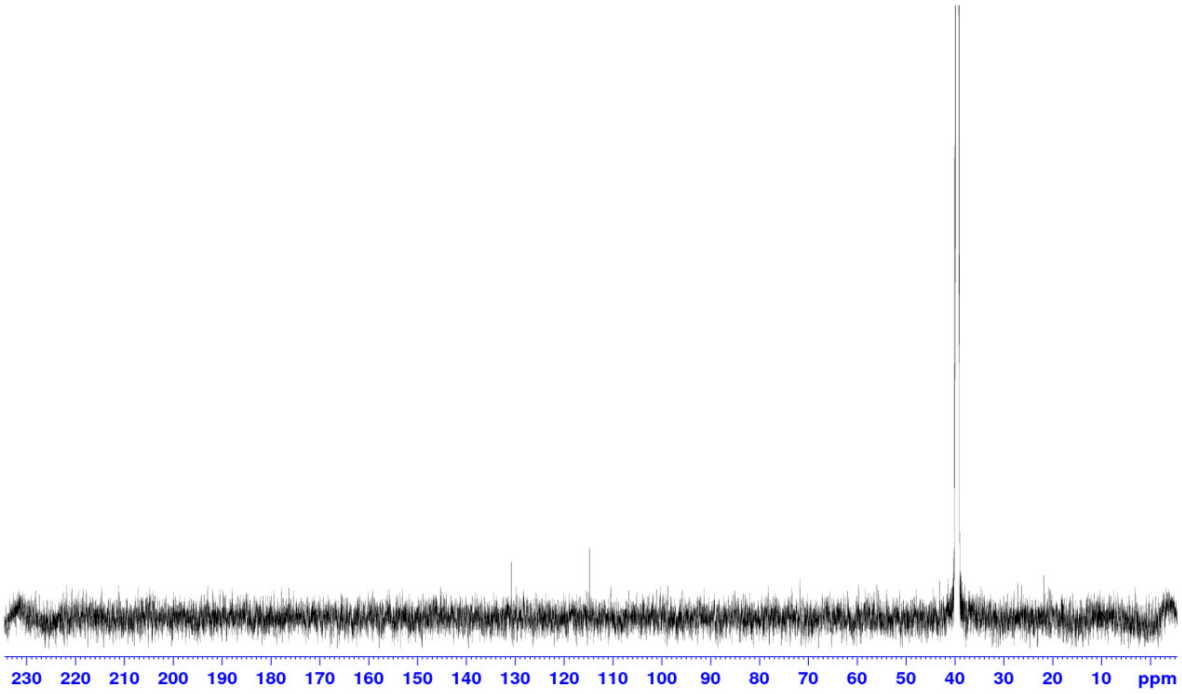
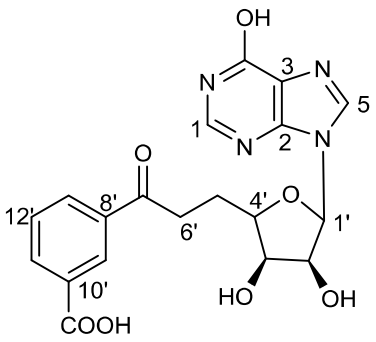


Figure S4. ^{13}C-NMR spectrum of (**2**) in DMSO-d_6.

futalosine (3): colorless oil: ^{1}H-NMR (600 MHz, DMSO-d_6) δ_H [ppm] (J[Hz]) 1.40 (2H, m, Ha-5'), 1.99 (2H, m, Hb-5'), 3.18 (2H, m, H-6'), 3.94 (1H, m, H-4'), 4.08 (1H, m, H-3'), 4.58 (1H, m, H-2'), 5.82 (1H, d, J 5.1, H-1'), 7.63 (1H, t, J 7.9, H-12'), 8.02 (1H, s, H-1), 8.14 (1H, d, J 7.9, H-11'), 8.16 (1H, d, J 7.9, H-13'), 8.28 (1H, s, H-5), 8.43 (1H, s, H-9'), 12.37 (1H, s, -COOH*). ^{13}C-NMR (600 MHz, DMSO-d_6) δ_C [ppm] 27.0 (C-5'), 34.3 (C-6'), 73.2 (C-3'), 74.2 (C-2'), 82.8 (C-4'), 88.1 (C-1'), 124.3 (C-3), 128.6 (C-12'), 130.6 (C-9'), 131.7 (C-10'), 132.5 (C-13'), 136.6 (C-11'), 136.7 (C-5), 146.5 (C-1), 147.9 (C-2), 156.5 (C-4), 166.6 (-*COOH), 198.9 (C-7').



3

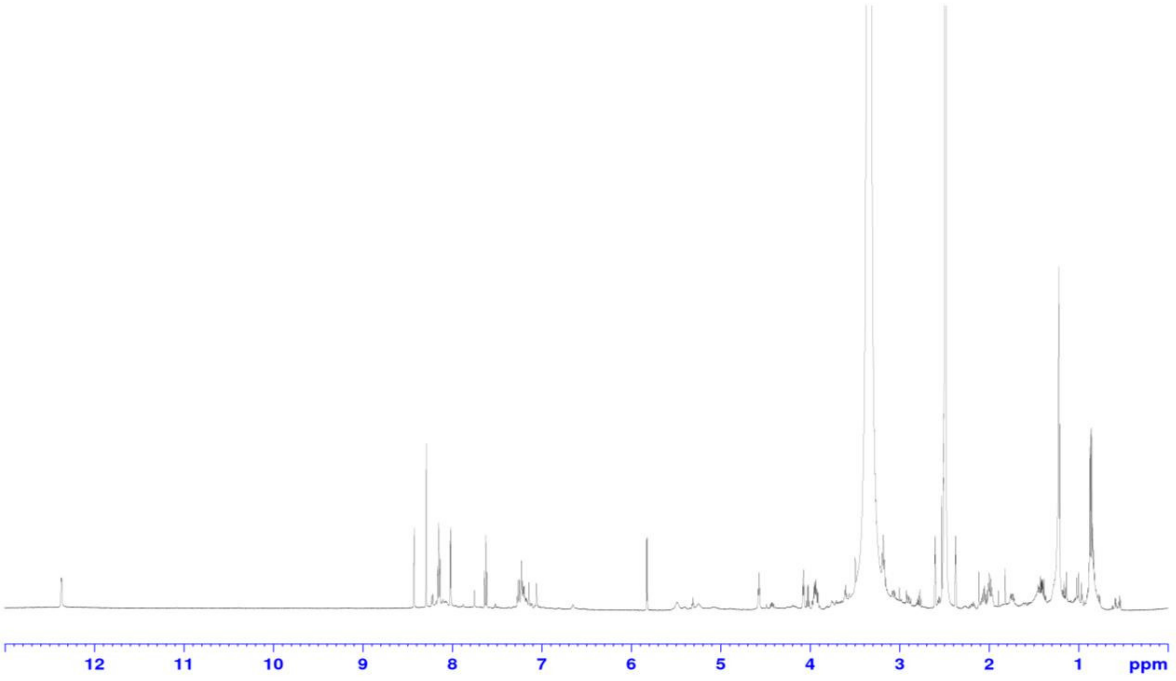


Figure S5. ^{1}H-NMR spectrum of (**3**) in DMSO-d_6.

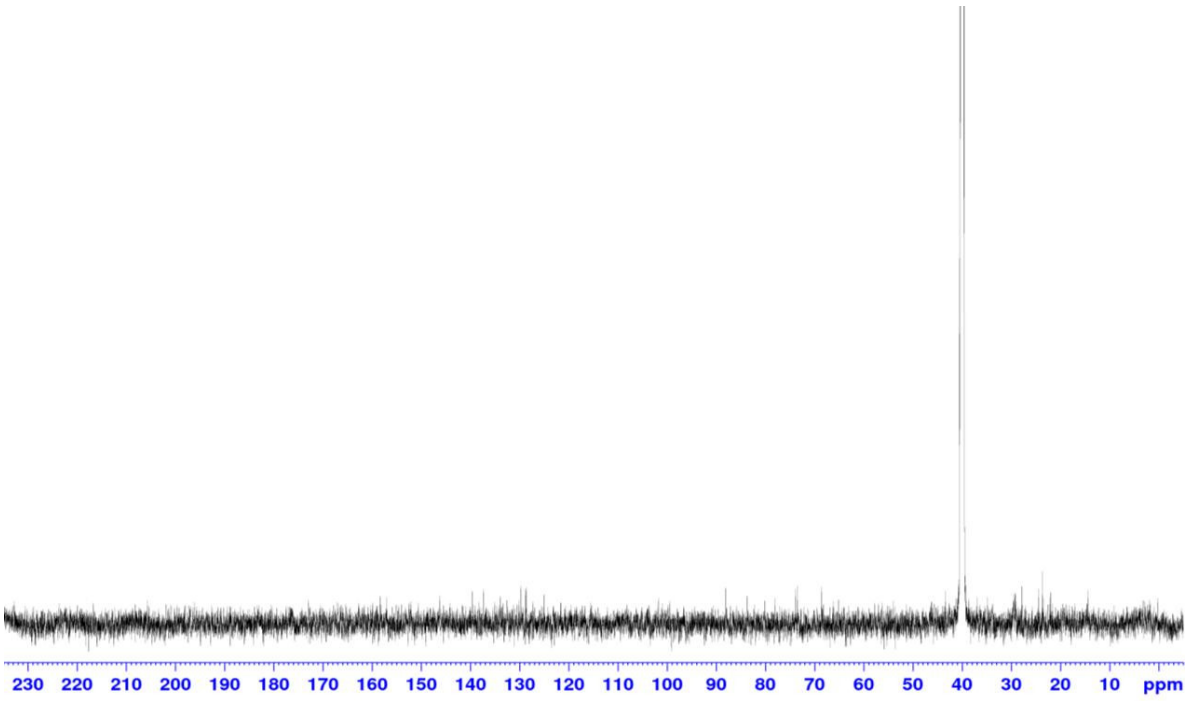
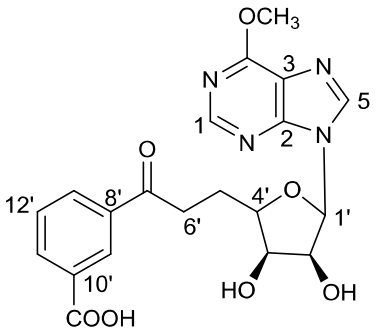


Figure S6. ^{13}C-NMR spectrum of (**3**) in DMSO-d_6.

4-*O*-methylfutalosine (4): colorless oil: ^{1}H-NMR (600 MHz, DMSO-d_6) δ_H [ppm] (*J*[Hz]) 2.15 (2H, m, H-5'), 2.74 (2H, m, Ha-6'), 3.02 (2H, m, Hb-6'), 3.33 (3H, s, -OCH$_3^*$), 3.91 (1H, m, H-4'), 4.23 (1H, t, *J* 5.5, H-3'), 4.94 (1H, dd, *J* 5.49, 5.56, H-2'), 5.92 (1H, d, *J* 7.6, H-1'), 7.72 (1H, m, H-12'), 7.96 (1H, d, *J* 7.7, H-13'), 8.03 (1H, s, H-1), 8.19 (1H, d, *J* 8.0, H-11'), 8.20 (1H, s, H-5), 8.38 (1H, s, H-9'). ^{13}C-NMR (600 MHz, DMSO-d_6) δ_C [ppm] 25.0 (C-6'), 30.1 (C-5'), 63.0 (-O*CH$_3$), 70.4 (C-2'), 76.0 (C-3'), 83.7 (C-4'), 86.4 (C-1'), 127.0 (C-9'), 127.0 (C-12'), 130.8 (C-3), 131.0 (C-13'), 134.1 (C-8'), 134.2 (C-10'), 136.2 (C-11'), 136.7 (C-5), 150.6 (C-2), 151.8 (C-1), 156.8 (C-4), 166.6 (-*COOH), 196.5 (C-7').



4

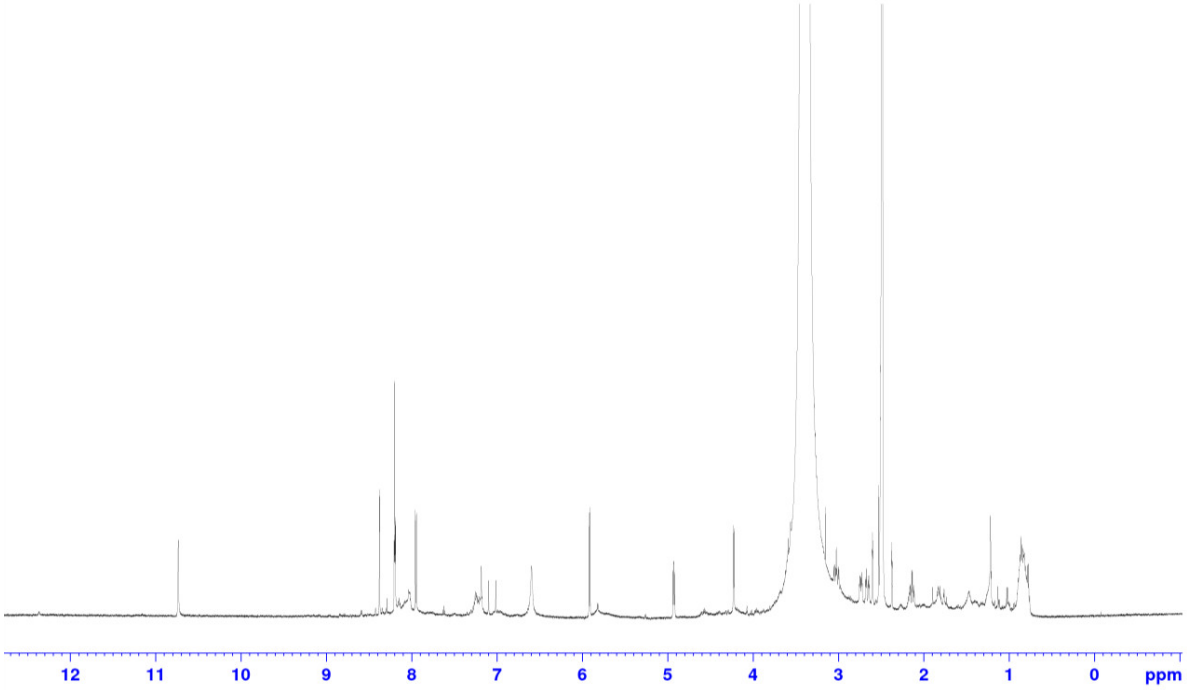


Figure S7. ^{1}H-NMR spectrum of (**4**) in DMSO-d_6.

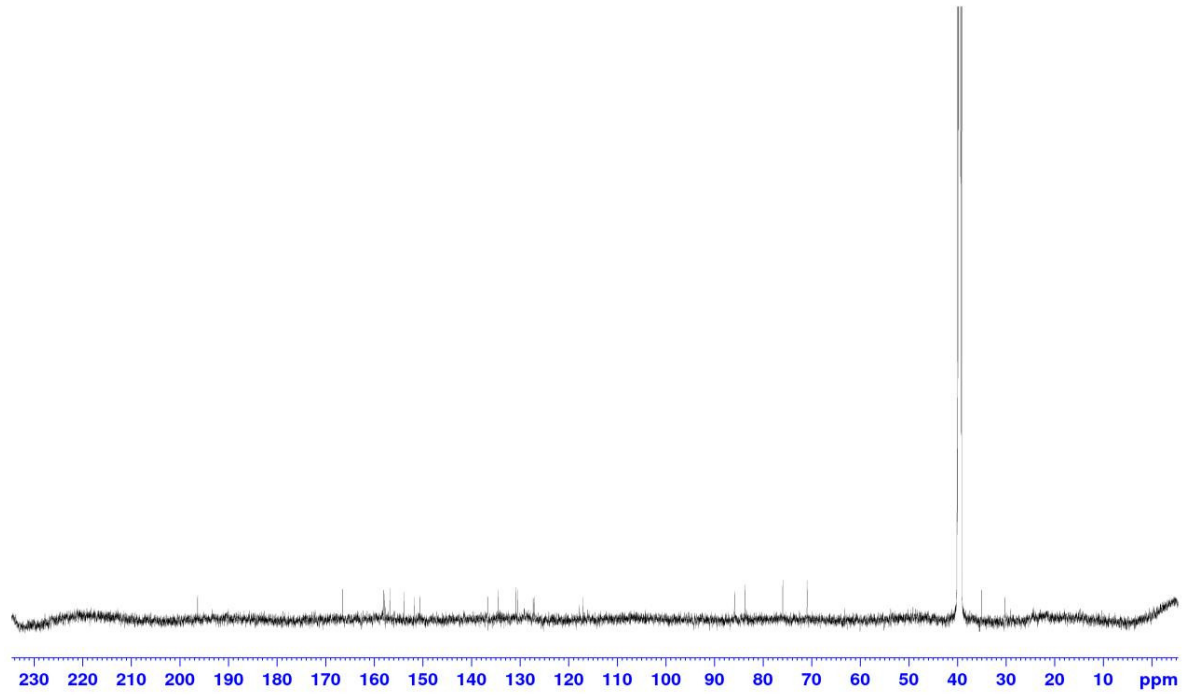
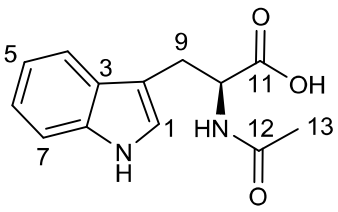


Figure S8. ^{13}C-NMR spectrum of (**4**) in DMSO-d_6.



5

***N*-acetyltryptophan (5):** colorless oil: ^{1}H-NMR (600 MHz, DMSO-d_6) δ_H [ppm] (*J*[Hz]) 1.79 (3H, s, H-13), 2.93 (2H, m, H-9), 4.47 (1H, m, H-10), 6.97 (1H, m, H-5), 7.05 (1H, m, H-6), 7.13 (1H, s, H-1), 7.33 (1H, d, *J* 8.2, H-7), 7.52 (1H, d, *J* 8.2, H-4), 8.13 (1H, s, -NH*), 10.82 (1H, s, -NH*). ^{13}C-NMR (600 MHz, DMSO-d_6) δ_C [ppm] 22.4 (C-13), 27.1 (C-9), 52.9 (C-10), 109.7 (C-2), 111.6 (C-7), 118.0 (C-4), 120.4 (C-5), 121.0 (C-6), 123.7 (C-1), 127.2 (C-3), 136.1 (C-8), 169.2 (C-12), 173.6 (C-11).

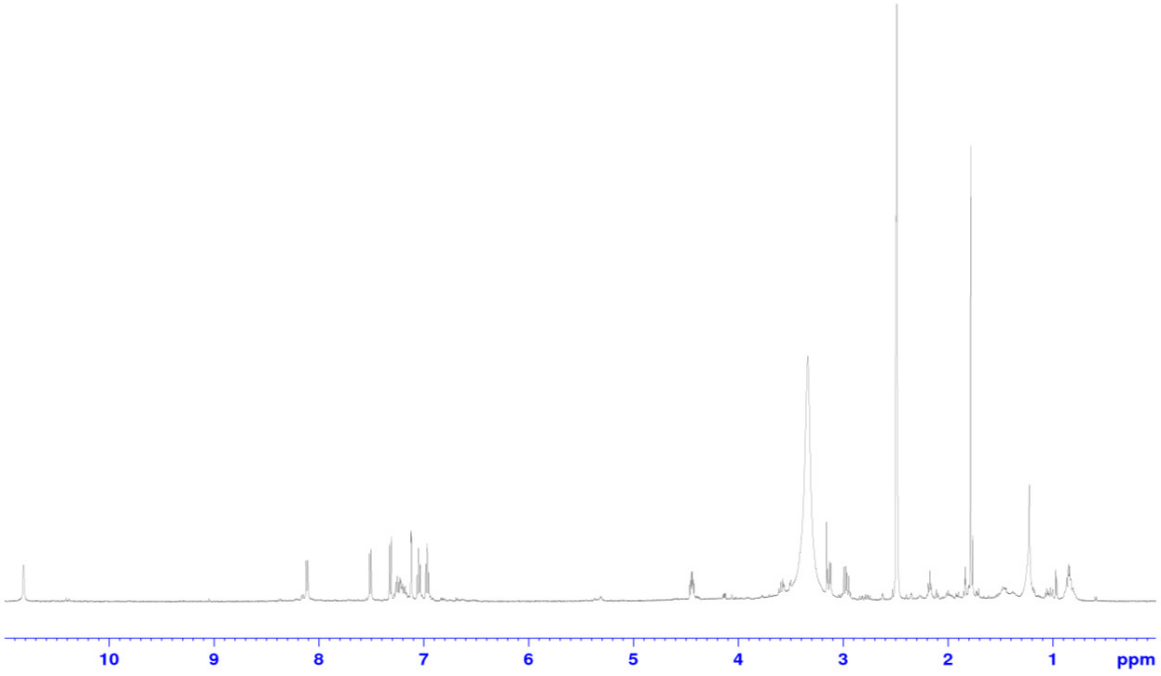


Figure S9. ^{1}H-NMR spectrum of (**5**) in DMSO-d_6.

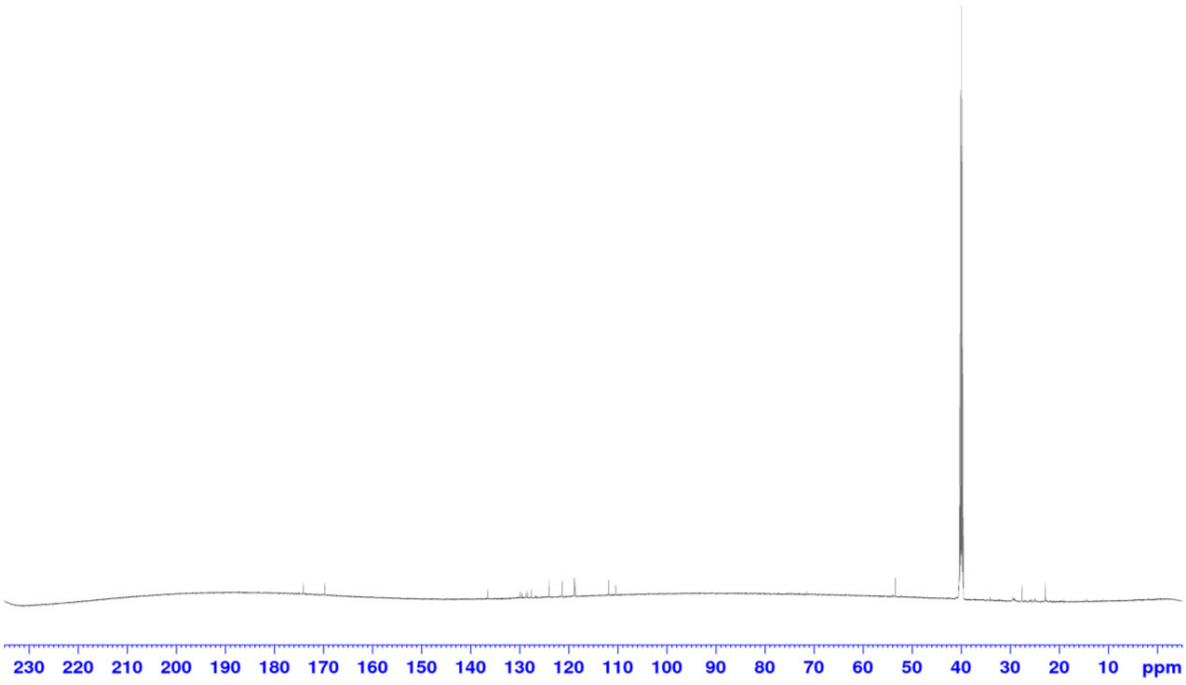
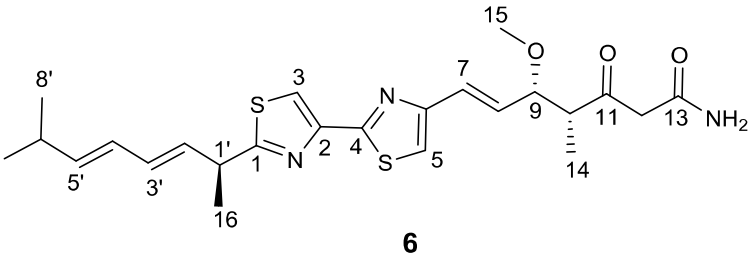


Figure S10. ^{13}C-NMR spectrum of (**5**) in DMSO-d_6.

desmethyl-myxothiazol (6): colorless oil: ^{1}H-NMR (600 MHz, DMSO-d_6) δ_H [ppm] (J[Hz]) 1.04 (3H, m, H-7', 8'), 1.14 (3H, d, J 7.1, H-14), 1.55 (3H, d, J 7.2, H-16), 2.35(1H, m, H-6'), 3.09 (1H, m, H-10), 3.35 (3H, s, H-15), 3.5 (2H, s, H-12), 3.94 (1H, m, H-1'), 4.11 (1H, dd, J 5.3, 7.3, H-9), 5.67 (1H, dd, J 5.7, 14.3, H-5'), 5.79 (1H, dd, J 7.9 15.7, H-2'), 6.05 (1H, dd, J 10.5, 15.3, H-4'),



6

6.23 (1H, dd, J 10.5, 15.3, H-3'), 6.45 (1H, dd, J 7.6, 15.3, H-8), 6.71 (1H, d, J 15.3, H-7), 7.43 (1H, s, H-

5), 8.04 (1H, s, H-3). ^{13}C-NMR (600 MHz, DMSO-d_6) δ_C [ppm] 10.6 (C-14), 19.6 (C-16), 21.2 (C-7', 8'), 30.8 (C-6'), 40.4 (C-12), 41 (C-1'), 50.6 (C-10), 55.4 (C-15), 82.6 (C-9), 115.9 (C-3), 116.5 (C-5), 126.4 (C-7), 126.7 (C-4'), 129.4 (C-8), 130.1 (C-2'), 131.0 (C-3'), 141.8 (C-5'), 148.3 (C-2), 153.7 (C-6), 176.9 (C-1), 177.9 (C-13), 206.4 (C-11).

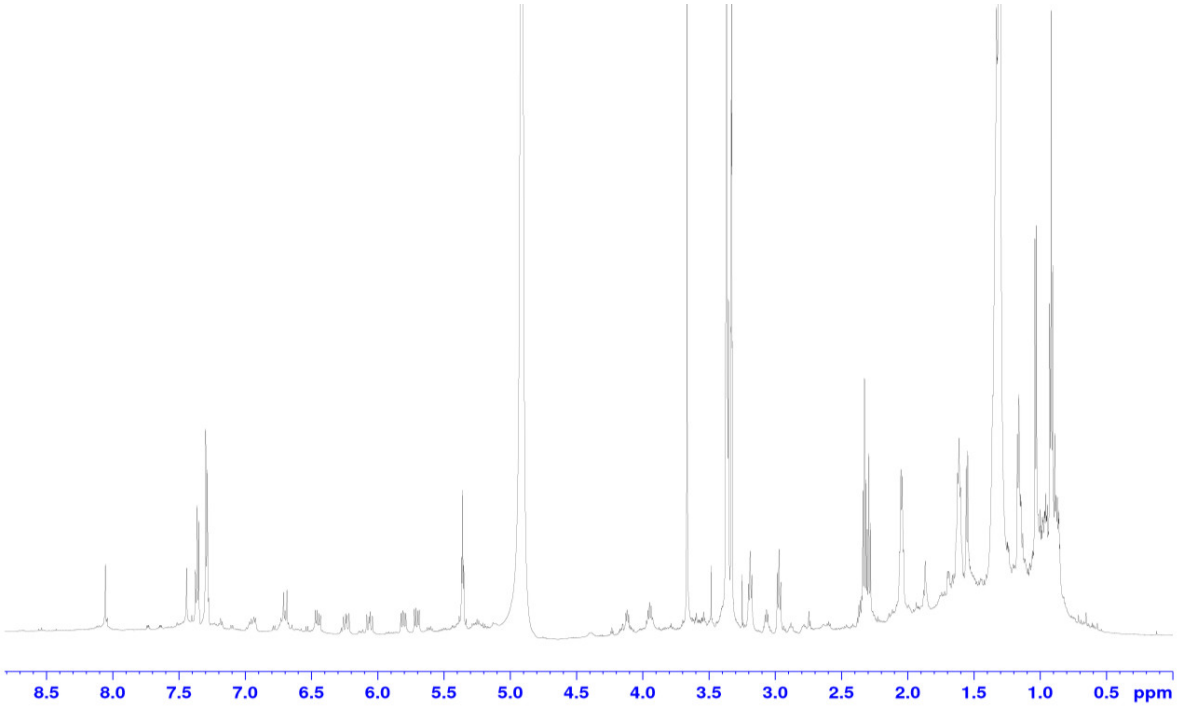


Figure S11. ^{1}H-NMR spectrum of (**6**) in DMSO-d_6.

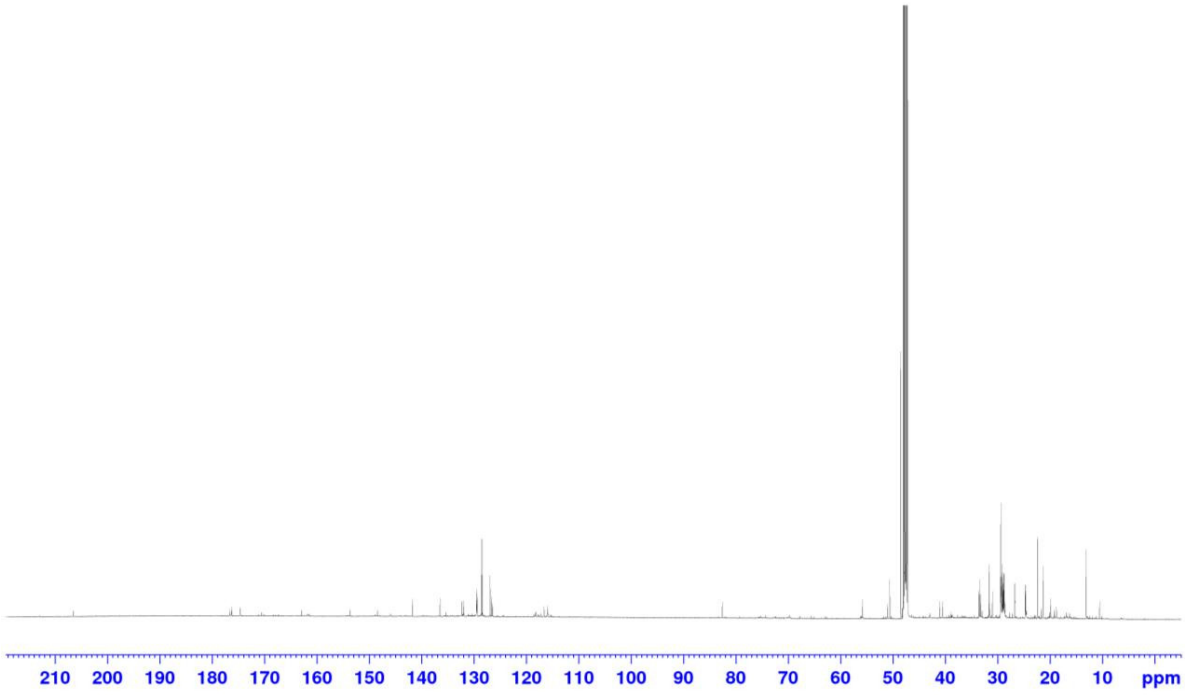


Figure S12. ^{13}C-NMR spectrum of (**6**) in DMSO-d_6.

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Die geltende Promotionsordnung der Biologisch-Pharmazeutischen Fakultät der Friedrich-Schiller-Universität Jena ist mir bekannt. Die Hilfe eines Promotionsberaters wurde nicht in Anspruch genommen. Es haben Dritte weder mittelbar noch unmittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit der vorgelegten Dissertation stehen. Diese schriftliche Arbeit wurde in gleicher oder ähnlicher Form noch bei keiner anderen Hochschule als Dissertation eingereicht und auch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung verwendet.

Jena, den _____

Xinli Pan

Abbreviations

NMR	nuclear magnetic resonance
HR-ESI-MS	high resolution electrospray ionization mass spectrometry
HPLC	high performance liquid chromatography
COSY	correlation spectroscopy
HMBC	heteronuclear multiple bond correlation
HSQC	heteronuclear single-quantum correlation
NOESY	nuclear overhauser effect spectroscopy
PCR	polymerase chain reaction
MK	menaquinone
BLAST	basic local alignment search tool
NRPS	nonribosomal peptide synthetase
PKS	polyketide synthase
Pfam	the protein families database
GC-content	guanine-cytosine content
DNA	deoxyribonucleic acid

Curriculum vitae

Personal Data

Name	Xinli Pan
Date of Birth	11.06.1987
Place of Birth	Nanning
Nationality	China

Education

Since 07/2016	Technical Biology, Department of Biochemical and Chemical Engineering, Technical University of Dortmund
Since 09/2013	Dissertation, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knöll-Institut, Junior research group "Natural Product Biosynthesis in the Genus *Herpetosiphon* – a Genomics and Chemistry-Guided Evaluation" Supervisor: Prof. Dr. Markus Nett and Prof. Dr. Christian Herweck
2010-2013	Pesticide Science, Huazhong Agricultural University Master Thesis: Isolation and identification of the secondary metabolites from *Coniothyrium minitans* Chy-1 and metabolomics analysis the different of *Rhizoctonia solani* AG-1-IA maturation and sclerotial. Supervisor: Dr. Xianwen Hu and Prof. Dr. Daohong Jiang
2005-2009	Applied chemistry, Huazhong Agricultural University Bachelor thesis: Investigation on the solubility of chitin Supervisor: Dr. Xianwen Hu

Publications

Pan XL, Domin N, Schieferdecker S, Kage H, Nett M. Herpetopanone, a diterpene from *Herpetosiphon aurantiacus* discovered by isotope labeling. *Beilstein J. Org. Chem*. **2017**, 13, 2458-2465.

Pan XL, Kage H, Martin K, Nett M. *Herpetosiphon gulosus* sp. nov., a filamentous predatory bacterium isolated from sandy soil and *Herpetosiphon giganteus* sp. nov., nom. rev.. *Int. J. Syst. Evol. Microbiol*. **2017**, 67, 2476-2481.

Pan XL and Nett M. Natural product biosynthesis in the phylum *Chloroflexi*: a genomics perspective. (Manuscript in preparation)

Research presentations

Pan XL, Domin N and Nett M. Diterpene Biosynthesis in *Herpetosiphon aurantiacu*s. ILRS symposium, 2014

Pan XL, Domin N and Nett M. Diterpene Biosynthesis in *Herpetosiphon aurantiacu*s. Micom, 2014

References

1. Cragg, G. M.; Newman, D. J., *Biochimi Biophys Acta.* **2013,** 1830 (6), 3670-3695.
2. Diggins, F. W., *Br. J. Biomed Sci.* **1999,** 56 (2), 83-93.
3. Lindner, E., Structure activities and pharmacological properties of the opium alkaloids. In *The chemistry and biology of isoquinoline alkaloids*, Springer: **1985**; pp 38-46.
4. Renslo, A. R., *ACS Med. Chem. Lett*. **2013**, 4(12), 1126-1128.
5. Fabbretti, A.; Gualerzi, C. O.; Brandi, L., *FEBS letters.* **2011,** 585 (11), 1673-1681.
6. Waksman, S. A.; Reilly, H. C.; Johnstone, D. B., *J Bacteriol.* **1946,** 52, 393-7.
7. Khokhlov, A. S.; Shutova, K. I., *J Antibiot (Tokyo).* **1972,** 25 (9), 501-8.
8. Ando, T.; Miyashiro, S.; Hirayama, K.; Kida, T.; Shibai, H.; Murai, A.; Udaka, S., *J Antibiot (Tokyo).* **1987,** 40 (8), 1140-5.
9. Basak, K.; Majumdar, S. K., *Antimicrob Agents Chemother.* **1973,** 4 (1), 6-10.
10. Chopra, I.; Roberts, M., *Microbiol. Mol. Biol. Rev.* **2001,** 65 (2), 232-260.
11. Kingston, D. G., *J. Nat. Prod.* **2010,** 74 (3), 496-511.
12. Naman, C. B.; Leber, C. A.; Gerwick, W. H., *Microbial Resources: From Functional Existence in Nature to Applications.* **2017**, 103.
13. Davies, S. C.; Fowler, T.; Watson, J.; Livermore, D. M.; Walker, D., *Lance.t* **2013,** 381 (9878), 1606-1609.
14. Andersson, D. I., *Curr. Opin. Microbiol.* **2003,** 6 (5), 452-456.
15. Gothwal, R.; Shashidhar, T., *Clean Soil Air Water.* **2015,** 43 (4), 479-489.
16. Schwartz, T.; Kohnen, W.; Jansen, B.; Obst, U., *FEMS Microbiol Ecol.* **2003,** 43 (3), 325-335.
17. Blair, J. M.; Webber, M. A.; Baylay, A. J.; Ogbolu, D. O.; Piddock, L. J., *Nat. Rev. Microbiol.* **2015,** 13 (1), 42-51.
18. Manchanda, V.; Sanchaita, S.; Singh, N., *J Glob Infec Dis.* **2010,** 2 (3), 291.
19. Hung, K.-H.; Wang, M.-C.; Huang, A.-H.; Yan, J.-J.; Wu, J.-J., *J. Clin. Microbiol.* **2012,** 50 (3), 721-726.
20. Gupta, N.; Limbago, B. M.; Patel, J. B.; Kallen, A. J., *Clin. Infec. Dis.* **2011,** 53 (1), 60-67.

21. Haddadin, A.; Fappiano, S.; Lipsett, P. A., *Postgrad. Med. J.* **2002,** 78 (921), 385-392.

22. Robert, J. G., *Clin. Infect. Dis.* **2011,** 52, 397.

23. Klevens, R. M.; Morrison, M. A.; Nadle, J.; Petit, S.; Gershman, K.; Ray, S.; Harrison, L. H.; Lynfield, R.; Dumyati, G.; Townes, J. M., *JAMA.* **2007,** 298 (15), 1763-1771.

24. Toutain, P. L.; Bousquet-Melou, A., *J. Vet. Pharmacol. Ther.* **2013,** 36 (5), 420-424.

25. Neu, H. C., *Science.* **1992,** 257 (5073), 1064-1074.

26. Elander, R., *App. Microbiol. Biotechnol.* **2003,** 61 (5-6), 385-392.

27. Birnbaum, J.; Kahan, F. M.; Kropp, H.; Macdonald, J. S., *Am. J. Med.* **1985,** 78 (6), 3-21.

28. Garrod, L. P., *Br. Med. J.* **1957,** 2 (5036), 57.

29. Tannock, G. W.; Munro, K.; Taylor, C.; Lawley, B.; Young, W.; Byrne, B.; Emery, J.; Louie, T., *Microbiology.* **2010,** 156 (11), 3354-3359.

30. Mullane, K. M.; Gorbach, S., *Expert Rev. Anti-Infect. Ther.* **2011,** 9 (7), 767-777.

31. Piscitelli, S.; Danziger, L.; Rodvold, K., *Clin. Pharm.* **1992,** 11 (2), 137-152.

32. Geddes, A. M.; Klugman, K. P.; Rolinson, G. N., *Int. J. Antimicrob. Agents.* **2007,** 30, 109-112.

33. Liang, D.; Wang, Y.; Wang, Y., *Mendeleev Communications.* **2015,** 25 (4), 252-253.

34. Kim, H. C.; Kang, S. H., *Angew. Chem. Int. Ed. Engl.* **2009,** 48 (10), 1827-1829.

35. Duran, D.; Aviyente, V.; Baysal, C., *J. Comput. Aided Mol. Des.* **2004,** 18 (2), 145-154.

36. Vieira, L. M. M.; de Almeida, M. V.; Lourenço, M. C. S.; Bezerra, F. A. F.; Fontes, A. P. S., *Eur. J. Med. Chem.* **2009,** 44 (10), 4107-4111.

37. Swaney, S. M.; Aoki, H.; Ganoza, M. C.; Shinabarger, D. L., *Antimicrob Agents Chemother.* **1998,** 42 (12), 3251-3255.

38. Lukacs, G.; Ohno, M., *Recent progress in the chemical synthesis of antibiotics*. Springer: New York, **2012**.

39. Berdy, J., *J. Antibiot (Tokyo).* **2005,** 58 (1), 1-26.

40. Winter, J. M.; Behnken, S.; Hertweck, C., *Curr. Opin. Chem. Biol.* **2011,** 15 (1), 22-31.

41. Reichenbach, H., *J. Ind. Microbiol Biotechnol.* **2001,** 27 (3), 149-156.

42. Burja, A. M.; Banaigs, B.; Abou-Mansour, E.; Burgess, J. G.; Wright, P. C., *Tetrahedron.* **2001,** 57 (46), 9347-9377.

43. Steinmetz, H.; Irschik, H.; Kunze, B.; Reichenbach, H.; Hoefle, G.; Jansen, R., *Chemistry.* **2007,** 13 (20), 5822-5832.

44. Simunovic, V.; Zapp, J.; Rachid, S.; Krug, D.; Meiser, P.; Müller, R., *Chembiochem.* **2006,** 7 (8), 1206-1220.

45. Chatterjee, S.; Chatterjee, S.; Lad, S. J.; Phansalkar, M. S.; Rupp, R.; Ganguli, B.; Fehlhaber, H.-W.; Kogler, H., *J. Antibiot (Tokyo).* **1992,** 45 (6), 832-838.

46. Milshteyn, A.; Schneider, J. S.; Brady, S. F., *Chem. Biol.* **2014,** 21 (9), 1211-1223.

47. Sockett, R. E.; Lambert, C., *Nat. Rev. Microbiol.* **2004,** 2 (8), 669-675.

48. Xiao, Y.; Wei, X.; Ebright, R.; Wall, D., *J. Bacteriol.* **2011,** 193 (18), 4626-4633.

49. Jurkevitch, E., *Microbe Am. Soc. Microbiol.* **2007,** 2 (2), 67.

50. Guerrero, R.; Pedrós-Alió, C.; Esteve, I.; Mas, J.; Chase, D.; Margulis, L., *Proc. Natl. Acad. Sci. U. S. A.* **1986,** 83 (7), 2138-2142.

51. Pérez, J.; Moraleda-Muñoz, A.; Marcos-Torres, F. J.; Muñoz-Dorado, J., *Environ Microbiol.* **2016,** 18 (3), 766-779.

52. Sockett, R. E., *Annu.l Rev. Microbiol.* **2009,** 63, 523-539.

53. Wang, Z.; Kadouri, D. E.; Wu, M., *BMC Genomics.* **2011,** 12 (1), 453.

54. Kumbhar, C.; Mudliar, P.; Bhatia, L.; Kshirsagar, A.; Watve, M., *Arch. Microbiol.* **2014,** 196 (4), 235-248.

55. Cavalier-Smith, T., *Int. J. Biochem. Cell. Biol.* **2009,** 41 (2), 307-322.

56. Casida, L., *Appl. Environ. Microbiol.* **1983,** 46 (4), 881-888.

57. Casida, L., *Microb Ecol.* **1988,** 15 (1), 1-8.

58. Konev, I.; Efimova, V.; Etingov, E.; Zaval'naia, N., *Antibiotiki.* **1978,** 23 (2), 143-148.

59. Imai, I.; Ishida, Y.; Hata, Y., *Mar. Biol.* **1993,** 116 (4), 527-532.

60. Banning, E. C.; Casciotti, K. L.; Kujawinski, E. B., *FEMS Microbiol Ecol.* **2010,** 73 (2), 254-270.

61. Lewin, R., *Microb Ecol.* **1997,** 34 (3), 232-236.

62. Holt, J.; Lewin, R., *J. Bacteriol.* **1968,** 95 (6), 2407.

63. Pan, X.; Kage, H.; Martin, K.; Nett, M., *Int. J. Syst. Evol. Microbiol.* **2017,** 67 (7), 2476-2481.

64. Guerrero, R.; Esteve, I.; PEDRÓS-ALIÓ, C.; Gaju, N., *Ann. NY. Acad. Sci.* **1987,** 503 (1), 238-250.

65. Grady, E. N.; MacDonald, J.; Liu, L.; Richman, A.; Yuan, Z.-C., *Microb. Cell Fact.* **2016,** 15 (1), 203.

66. Raza, W.; Yang, W.; Shen, Q., *Plant Pathol.* **2008**, 419-430.

67. Beatty, P. H.; Jensen, S. E., *Can. J. Microbiol.* **2002,** 48 (2), 159-169.

68. Willems, A.; Fernández-López, M.; Munoz-Adelantado, E.; Goris, J.; De Vos, P.; Martínez-Romero, E.; Toro, N.; Gillis, M., *Int. J. Syst. Evol. Microbiol.* **2003,** 53 (4), 1207-1217.

69. Kuever, J.; Rainey, F.; Widdel, F., *Desulfobulbus*. In: Brenner DJKNaSJ, editor. Bergey's Manual of Systematic Bacteriology Volume 2: The Proteobacteria; Part C The Alpha-, Beta-, Delta and Epsilonproteobacteria. New York: Springer. pp. 988–992.J. KueverFA RaineyF. Widdel, 2005. *Desulfobulbus*. DJKNaSJ BrennerBergey's Manual of Systematic Bacteriology Volume 2: The Proteobacteria; Part C The Alpha-, Beta-, Delta and Epsilonproteobacteria New York Springer 988992. **2005**.

70. Seccareccia, I.; Kovács, Á. T.; Gallegos-Monterrosa, R.; Nett, M., *Microbiol Res.* **2016,** 192, 231-238.

71. Makkar, N.; Casida Jr, L., *Int. J. Syst. Evol. Microbiol.* **1987,** 37 (4), 323-326.

72. Seccareccia, I.; Kost, C.; Nett, M., *Appl. Environ. Microbiol.* **2015,** 81 (20), 7098-7105.

73. Kaparullina, E.; Doronina, N.; Chistyakova, T.; Trotsenko, Y., *Syst. Appl. Microbiol.* **2009,** 32 (3), 157-162.

74. Koval, S. F.; Williams, H. N.; Stine, O. C., *Int. J. Syst. Evol. Microbiol.* **2015,** 65 (2), 593-597.

75. Berleman, J. E.; Kirby, J. R., *FEMS microbiol Rev.* **2009,** 33 (5), 942-957.

76. Davidov, Y.; Jurkevitch, E., *Int. J. Syst. Evol. Microbiol.* **2004,** *54* (5), 1439-1452.

77. Kimura, Y.; Mori, Y.; Ina, Y.; Takegawa, K., *J. Bacteriol.* **2011,** 193 (10), 2657-2661.

78. Evans, A. G.; Davey, H. M.; Cookson, A.; Currinn, H.; Cooke-Fox, G.; Stanczyk, P. J.; Whitworth, D. E., *Microbiology.* **2012,** 158 (11), 2742-2752.

79. Livingstone, P. G.; Morphew, R. M.; Whitworth, D. E., *Front Microbiol.* **2017,** 8, 1593.

80. Lambert, C.; Chang, C.-Y.; Capeness, M. J.; Sockett, R. E., *PloS one.* **2010,** 5 (1), e8599.

81. Korp, J.; Gurovic, M. S. V.; Nett, M., *Beilstein J. Org. Chem.* **2016,** 12, 594.

82. Iebba, V.; Totino, V.; Santangelo, F.; Gagliardi, A.; Ciotoli, L.; Virga, A.; Ambrosi, C.; Pompili, M.; De Biase, R. V.; Selan, L., *Front Microbiol.* **2014,** 5.

83. Dashiff, A.; Junka, R. A.; Libera, M.; Kadouri, D. E., *J. Appl. Microbiol.* **2011,** 110 (2), 431-44.

84. Fothergill, J. L.; Winstanley, C.; James, C. E., *Expert. Rev. Anti-Infect. Ther.* **2012,** 10 (2), 219-235.

85. Withey, S.; Cartmell, E.; Avery, L.; Stephenson, T., *Sci. Total Environ.* **2005,** 339 (1), 1-18.

86. Garrity, G. M. Holt, J. G., Phylum BVI. *Chlorofexi* phy. nov. In *Bergey's Manual of Systematic Bacteriology*, 2nd edn, vol. 1 (*The Archaea and the Deeply Branching and Phototrophic Bacteria*), pp. 427-446. Edited by D. R. Boone, R. W. Castenholz & G. M. Garrity. New York: Springer. **2001**.

87. Lewin, R. A.; Leadbetter. R. E., Genus *Herpetosiphon* Holt and Lewin 1968, 2408; emend. mut. char. In: Buchanan RE, Gibbons NE (eds), *Bergey's Manual of Determinative Bacteriology*, Eighth Edition, The Williams and Wilkins Co., Baltimore, **1974**, 107-109.

88. Jürgen, U.; Meissner, J.; Reichenbach, H.; Weckesser, J., *FEMS Microbiol Lett.* **1989,** 60 (3), 247-250.

89. Kiss, H.; Nett, M.; Domin, N.; Martin, K.; Maresca, J. A.; Copeland, A.; Lapidus, A.; Lucas, S.; Berry, K. W.; Del Rio, T. G., *Stand Genomic Sci.* **2011,** 5 (3), 356.

90. Ward, L. M.; Hemp, J.; Pace, L. A.; Fischer, W. W., *Genome Announc.* **2015,** 3 (6), e01352-15.

91. Lewin, R. A., *Can. J. Microbiol.* **1970,** 16 (6), 517-520.

92. Nett, M.; Erol, Ö.; Kehraus, S.; Köck, M.; Krick, A.; Eguereva, E.; Neu, E.; König, G. M., *Angew. Chem. Int. Ed. Engl.* **2006,** 45 (23), 3863-3867.

93. Mir Mohseni, M.; Höver, T.; Barra, L.; Kaiser, M.; Dorrestein, P. C.; Dickschat, J. S.; Schäberle, T. F., *Angew. Chem. Int. Ed. Engl.* **2016,** 55 (43), 13611-13614.

94. Schieferdecker, S.; Domin, N.; Hoffmeier, C.; Bryant, D. A.; Roth, M.; Nett, M., *European J. Org. Chem.* **2015,** 2015 (14), 3057-3062.

95. Nakano, C.; Oshima, M.; Kurashima, N.; Hoshino, T., *Chembiochem.* **2015,** 16 (5), 772-781.

96. Kodama, M.; Okumura, K.; Kobayashi, Y.; Tsunoda, T.; Itô, S., *Tetrahedron lett.* **1984,** 25 (50), 5781-5784.

97. Yamada, Y.; Kuzuyama, T.; Komatsu, M.; Shin-ya, K.; Omura, S.; Cane, D. E.; Ikeda, H., *Proc. Natl. Acad. Sci. U. S. A.* **2015,** 112 (3), 857-862.

98. Andersen, N.; Shunk, B.; Costin, C., *Cell. Mol. Life Sci.* **1973,** 29 (6), 645-647.

99. Wang, G.; Li, X.; Huang, F.; Zhao, J.; Ding, H.; Cunningham, C.; Coad, J.; Flynn, D.; Reed, E.; Li, Q., *Cell. Mol. Life Sci.* **2005,** *62* (7), 881-893.

100. Citron, C. A.; Rabe, P.; Dickschat, J. S., *J. Nat. Prod. Rep.* **2012,** 75 (10), 1765-1776.

101. Zhang, J.; Polishchuk, E. A.; Chen, J.; Ciufolini, M. A., *J. Org. Chem.* **2009,** 74 (23), 9140-9151.

102. Nikolouli, K.; Mossialos, D., *Biotechnology letters* **2012,** *34* (8), 1393-1403.

103. Hertweck, C., *Angew. Chem. Int. Ed.* **2009,** 48 (26), 4688-4716.

104. Staunton, J.; Weissman, K. J., *Nat. Prod. Rep.* **2001,** 18 (4), 380-416.

105. Hertweck, C.; Luzhetskyy, A.; Rebets, Y.; Bechthold, A., *Nat. Prod. Rep.* **2007,** *24* (1), 162-190.

106. Yan, X.; Probst, K.; Linnenbrink, A.; Arnold, M.; Paululat, T.; Zeeck, A.; Bechthold, A., *Chembiochem.* **2012,** 13 (2), 224-230.

107. Oki, T.; Tenmyo, O.; Hirano, M.; Tomatsu, K.; Kamel, H., *J. Antibiot (Tokyo).* **1990,** 43 (7), 763-770.

108. Fitzgerald, J. T.; Henrich, P. P.; O'brien, C.; Krause, M.; Ekland, E. H.; Mattheis, C.; Sá, J. M.; Fidock, D.; Khosla, C., *J. Antibiot (Tokyo).* **2011,** 64 (12), 799.

109. Li, Y.; Müller, R., *Phytochemistry.* **2009,** 70 (15), 1850-1857.

110. Katsuyama, Y.; Ohnishi, Y., *Methods Enzymol.* **2012,** 515, 359-377.

111. Nakano, C.; Ozawa, H.; Akanuma, G.; Funa, N.; Horinouchi, S., *J. Bacteriol.* **2009,** 191 (15), 4916-4923.

112. Miao, V.; Coeffet-LeGal, M.-F.; Brian, P.; Brost, R.; Penn, J.; Whiting, A.; Martin, S.; Ford, R.; Parr, I.; Bouchard, M., *Microbiology.* **2005,** 151 (5), 1507-1523.

113. Moyne, A.-L.; Cleveland, T. E.; Tuzun, S., *FEMS Microbiol Lett.* **2004,** 234 (1), 43-49.

114. Kunze, B.; Bedorf, N.; Kohl, W.; Höfle, G.; Reichenbach, H., *J. Antibiot (Tokyo).* **1989,** 42 (1), 14-17.

115. Silakowski, B.; Kunze, B.; Nordsiek, G.; Blöcker, H.; Höfle, G.; Müller, R., *FEBS J.* **2000,** 267 (21), 6476-6485.

116. Canafax, D.; Ascher, N., *Clin. Pharm.* **1983,** 2 (6), 515-524.

117. Pohle, S.; Appelt, C.; Roux, M.; Fiedler, H.-P.; Süssmuth, R. D., *J. Am. Chem. Soc.* **2011,** 133 (16), 6194-6205.

118. Schneider, A.; Marahiel, M. A., *Arch. Microbiol.* **1998,** 169 (5), 404-410.

119. Stachelhaus, T.; Marahiel, M. A., *FEMS microbiol Lett.* **1995,** 125 (1), 3-14.

120. Keating, T. A.; Marshall, C. G.; Walsh, C. T., *Biochemistry.* **2000,** 39 (50), 15513-15521.

121. Pfeifer, B. A.; Wang, C. C.; Walsh, C. T.; Khosla, C., *Appl. Environ. Microbiol.* **2003,** 69 (11), 6698-6702.

122. Lange, B. M.; Rujan, T.; Martin, W.; Croteau, R., *Proc. Natl. Acad. Sci. U. S. A.* **2000,** 97 (24), 13172-13177.

123. Heuston, S.; Begley, M.; Gahan, C. G.; Hill, C., *Microbiology.* **2012,** 158 (6), 1389-1401.

124. Christianson, D. W., *Chem. Rev.* **2017,** 117 (17), 11570-11648.

125. Dickschat, J. S., *Nat. Prod. Rep.* **2016,** 33 (1), 87-110.

126. Thulasiram, H. V.; Erickson, H. K.; Poulter, C. D., *J. Am. Chem. Soc.* **2008,** 130 (6), 1966-1971.

127. Mandal, S. C.; Mandal, V.; Das, A. K., *Essentials of botanical extraction: principles and applications*. Academic Press, Elseiver, London, 1st edn, **2015**.

128. Ziemert, N.; Alanjary, M.; Weber, T., *Nat. Prod. Rep.* **2016,** 33 (8), 988-1005.

129. Zerikly, M.; Challis, G. L., *Chembiochem.* **2009,** 10 (4), 625-633.

130. Katz, L.; Baltz, R. H., *J. Ind. Microbiol.* **2016,** 43 (2-3), 155-176.

131. Bachmann, B. O.; Van Lanen, S. G.; Baltz, R. H., *J. Intd. Microbiol.* **2014,** 41 (2), 175-184.

132. van Heel, A. J.; de Jong, A.; Montalban-Lopez, M.; Kok, J.; Kuipers, O. P., *Nucleic Acids Res.* **2013,** 41 (1), 448-453.

133. Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Bruccoleri, R.; Lee, S. Y.; Fischbach, M. A.; Müller, R.; Wohlleben, W., *Nucleic Acids Res.* **2015,** 43 (1), 237-243.

134. Agrawal, P.; Khater, S.; Gupta, M.; Sain, N.; Mohanty, D., *Nucleic Acids Res.* **2017,** 45(1), 80-88.

135. Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A.-M.; Challis, G. L.; Thomson, N.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D., *Nature.* **2002,** 417 (6885), 141-147.

136. Nett, M.; Ikeda, H.; Moore, B. S., *Nat. Prod. Rep.* **2009,** 26 (11), 1362-1384.

137. Amoutzias, G. D.; Chaliotis, A.; Mossialos, D., *Mar. Drugs.* **2016,** 14 (4), 80.

138. Ikeda, H.; Shin-ya, K.; Omura, S., *J. Ind. Microbiol.* **2014,** 41 (2), 233-250.

139. Machado, H.; Sonnenschein, E. C.; Melchiorsen, J.; Gram, L., *BMC Genomics.* **2015,** 16 (1), 158.

140. Gomez-Escribano, J. P.; Bibb, M. J., *J. Ind. Microbiol.* **2014,** 41 (2), 425-431.

141. Nah, H.-J.; Pyeon, H.-R.; Kang, S.-H.; Choi, S.-S.; Kim, E.-S., *Front Microbiol.* **2017,** 8, 394.

142. Hewage, R. T.; Aree, T.; Mahidol, C.; Ruchirawat, S.; Kittakoop, P., *Phytochemistry.* **2014,** 108, 87-

94.

143. Bode, H. B.; Bethe, B.; Höfs, R.; Zeeck, A., *Chembiochem.* **2002,** 3 (7), 619-627.

144. Hussain, A.; Rather, M. A.; Dar, M. S.; Aga, M. A.; Ahmad, N.; Manzoor, A.; Qayum, A.; Shah, A.; Mushtaq, S.; Ahmad, Z., *Bioorganic Med. Chem. Lett.* **2017,** 27 (11), 2579-2582.

145. Gross, H.; Stockwell, V. O.; Henkels, M. D.; Nowak-Thompson, B.; Loper, J. E.; Gerwick, W. H., *Chem. Biol.* **2007,** 14 (1), 53-63.

146. Hosokawa, N.; Naganawa, H.; Kasahara, T.; Hattori, S.; Hamada, M.; Takeuchi, T.; Yamamoto, S.; Tsuchiya, K. S.; Hori, M., *Chem. Pharm. Bull.* **1999,** 47 (7), 1032-1034.

147. Schieferdecker, S.; Exner, T. E.; Gross, H.; Roth, M.; Nett, M., *J. Antibiotic (Tokyo).* **2014,** 67 (7), 519-525.

148. Böhlendorf, B.; Herrmann, M.; Hecht, H. J.; Sasse, F.; Forche, E.; Kunze, B.; Reichenbach, H.; Höfle, G., *Eur. J. Org. Chem.* **1999,** 1999 (10), 2601-2608.

149. Hefter, J.; Richnow, H. H.; Fischer, U.; Trendel, J. M.; Michaelis, W., *Microbiology.* **1993,** 139 (11), 2757-2761.

150. Pan, X.; Domin, N.; Schieferdecker, S.; Kage, H.; Roth, M.; Nett, M., *Beilstein J. Org. Chem.* **2017,** *13*, 2458.

151. Boronat, A.; Rodríguez-Concepción, M., Terpenoid biosynthesis in prokaryotes. *Adv Biochem Eng Biotechnol*. **2015**, 148, 3-18.

152. Marco, J. A.; Sanz-Cervera, J. F.; Sancenon, F.; Jakupovic, J.; Rustaiyant, A.; Mohamadit, F., *Phytochemistry.* **1993,** 34 (4), 1061-1065.

153. Wong, K.; Hamid, A.; Eldeen, I.; Asmawi, M. Z.; Baharuddin, S.; Abdillahi, H.; Staden, J. V., *Nat. Prod. Res.* **2012,** 26 (9), 850-858.

154. Tchuendem, M. H.; Mbah, J. A.; Tsopmo, A.; Ayafor, J. F.; Sterner, O.; Okunjic, C. C.; Iwu, M. M.; Schuster, B. M., *Phytochemistry.* **1999,** 52 (6), 1095-1099.

155. Chen, P.; Wang, P. P.; Jiao, Z. Z.; Xiang, L., *Helv. Chim. Acta.* **2014,** 97 (3), 388-397.

156. Miyanaga, S.; Obata, T.; Onaka, H.; Fujita, T.; Saito, N.; Sakurai, H.; Saiki, I.; Furumai, T.; Igarashi, Y., *J. Antibiot (Tokyo).* **2006,** 59 (11), 698-703.

157. Jakobi, K.; Hertweck, C., *J. Am. Chem. Soc.* **2004,** 126 (8), 2298-2299.

158. Meyer, J.-M.; Geoffroy, V. A.; Baida, N.; Gardan, L.; Izard, D.; Lemanceau, P.; Achouak, W.; Palleroni, N. J., *Appl. Environ. Microbiol.* **2002,** 68 (6), 2745-2753.

159. Foschi, C.; Laghi, L.; Parolin, C.; Giordani, B.; Compri, M.; Cevenini, R.; Marangoni, A.; Vitali, B., *PloS one.* **2017,** 12 (2), e0172483.